

MDS and AML associated DDX41 helicase resolves R-loops to maintain genome stability

A Thesis Submitted to the College of
Graduate and Postdoctoral Studies
In Fulfillment of the
Requirements for the Degree of
Master of Science
In the Department of Biochemistry, Microbiology, and
Immunology
University of Saskatchewan
Saskatoon

By

Aanchal Aggarwal

© Copyright Aanchal Aggarwal, December 2021. All rights reserved
Unless otherwise noted, copyright of the material in this thesis belongs to the author

PERMISSION OF USE STATEMENT

I hereby present this thesis in partial fulfilment of the requirements for a postgraduate degree from the University of Saskatchewan and agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis in any manner, either in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised this thesis or, in their absence, by the Head of the Department or the Dean of the College in which my thesis work was done. It is understood that any copying or publication or use of this thesis or parts of it for any financial gain will not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis.

Requests for permission to copy or to make other use of material in this thesis in whole or part should be addressed to:

Dr. Bill Roesler

Head of the Department of Biochemistry, Microbiology, and Immunology

University of Saskatchewan

Saskatoon, Saskatchewan, S7N 5E5

Dr. Debby Burshtyn

Dean

College of Graduate and Postdoctoral Studies

University of Saskatchewan

116 Thorvaldson Building, 110 Science Place

Saskatoon, Saskatchewan, S7N 5C9

Canada

ABSTRACT

DNA damage and aberrant DNA damage response (DDR) are significant features of genomic instability that are implicated in the pathogenesis of myelodysplastic syndromes (MDS) and acute myeloid leukemia (AML). DEAD-box helicase 41 (DDX41) is a member of helicase superfamily 2. Both germline and somatic mutations of DDX41 have been associated with MDS and AML, and missense mutant R525H is the most frequent mutation (67%) found in MDS and AML patients. However, the molecular pathogenesis of DDX41 mutations, such as R525H, remains unknown. Here, I found that DDX41 protein could efficiently unwind RNA:DNA hybrid *in vitro*. This prompted us to investigate the potential role of DDX41 in DDR concerning R-loops, one of the significant sources of genome instability. Examining DDX41 knockout (KO) cells (HeLa and HT1080) and DDX41 KO cells reconstituted with *DDX41-R525H* gene, I found that these cells had prolonged and increased DNA double-strand breaks (DSB) induced by ionizing radiation (IR) or bleomycin, a radiomimetic agent, compared to wildtype (WT) and DDX41 KO cells reconstituted with DDX41-WT gene, suggesting that DDX41 is crucial for DDR. Using the R-loop specific antibody S9.6, dot blot assays showed increased R-loops in DDX41-KO and R525H-expressing cells, deciphering that DDX41 functions as an R-loop resolvase. Furthermore, biochemical assays demonstrated R525H stimulates R-loops formation. Lastly, I found that DDX41 colocalized with the DSB marker γ H2AX, R-loop marker S9.6, and transcription machinery marker RNA polymerase II, indicating that DDX41 interacts with proteins during transcription and DNA repair. In conclusion, our results imply that DDX41 is required for DDR as an R-loop resolvase upon DNA damage, and lack of or impaired DDX41 results in altered DDR response and pathogenic R-loop accumulation, which ultimately leads to genome instability and possibly pathogenesis of MDS and AML.

ACKNOWLEDGEMENTS

I want to express my deepest gratitude to my supervisor Dr. Yuliang Wu for giving me the opportunity to carry out this project in his lab, his endless support on academic studies and his helpful guidance on my career. It is his level of patience, knowledge, and ingenuity that I will always keep aspiring to. I would also like to thank my committee members Dr. Scot Leary, Dr. Jeremy Lee, and Dr. Yan Zhou, for their valuable inputs, insightful suggestions, and guidance throughout my project. Together, my work results from constant encouragement and ideas from this team of professors, and I am obliged to work under their supervision. I am also thankful to all Dr. Wu's lab members who helped me find my footing as I started this process. Dr. Ravi Shankar Singh, Dr. Venkatasubramanian Vidhyasagar, Shizhuo (Sarah) Yang, Manisha Yadav, Kingsley Ekumi, Ananna Arna, He Dong all deserve appreciation for their persistent guidance and support. I am fortunate to have been a part of the Cancer Cluster, and I would like to acknowledge all members in the Department of Biochemistry, Microbiology, and Immunology. My gratitude and appreciation go to our lab manager, Mark Boyd for always being there as a 'human trouble-shooter' throughout the program. It is important to strike a balance with life outside the depths of the lab. I cannot stress enough the importance of my get-togethers with Nayoung Kim, Dr. Amrutha Balagopal, Dr. Karthic Rajamanickam, Hardikkumar Patel, Stephany Cornea, Aditya Mandapati, and Hussain Elhasasna.

I must express my profound gratitude to my parents and brother for providing me with continuous encouragement to pursue my dreams and for their patience throughout my years of research. I am also thankful to my best friend, Swati Mishra and all my other friends for their help and motivation throughout my program.

Thank you to my partner, Ujjaval Patel, for constantly listening to my rants and staying by my side, for feeding me delicious food, for all the trips, and for cracking jokes during hard days, and for the sacrifices he has made for me to pursue this program.

TABLE OF CONTENTS

PERMISSION OF USE STATEMENT.....	ii
ABSTRACT.....	iii
ACKNOWLEDGEMENTS.....	iv
TABLE OF CONTENTS.....	v
LIST OF FIGURES.....	vii
LIST OF ABBREVIATIONS.....	ix
1. INTRODUCTION.....	1
1.1 Helicases.....	1
1.2 DEAD-box helicases.....	2
1.3 DDX41 helicase.....	4
1.4 DDX41 mutations in MDS/AML.....	5
1.5 DNA damage and repair.....	7
1.6 R-loops.....	10
2. HYPOTHESIS AND OBJECTIVES.....	13
2.1 Hypothesis.....	13
2.2 Objectives.....	13
3. MATERIALS AND METHODS.....	14
3.1 Plasmid DNA.....	14
3.2 Cell lines.....	14
3.3 CRISPR knockout.....	15
3.4 Reagents.....	15
3.5 DNA damage treatments.....	17
3.6 Antibodies.....	17
3.7 Confocal microscopy.....	18
3.8 Western blotting.....	18
3.9 Alkaline comet assays.....	19

3.10 Dot blot assays.....	19
3.11 Recombinant proteins.....	20
3.12 RNA and DNA substrates.....	21
3.13 Helicase assays.....	22
3.14 Strand annealing assays.....	22
4. RESULTS.....	22
4.1 DDX41 forms foci upon IR exposure.....	22
4.2 DDX41 is involved in the DNA damage response pathway.....	24
4.3 Increased DNA damage in DDX41-knockout cells.....	26
4.4 DDX41 co-localizes with the DNA double-strand break marker γ H2AX.....	29
4.5 Comet assay validates higher and prolonged DNA damage in DDX41-KO cells.....	32
4.6 DDX41 promotes R-loop clearance.....	33
4.7 Purification of DDX41-WT and DDX41-R525H mutant proteins.....	34
4.8 R525H protein has reduced unwinding activity but retains normal strand annealing activity.....	35
4.9 DDX41 co-localizes with R-loops.....	37
4.10 DDX41 co-localizes with transcriptional machinery.....	38
4.11 Patient mutant R525H exhibits diminished DDR and increased DNA damage when reconstituted in wildtype or DDX41 KO cells.....	40
5. DISCUSSION.....	44
5.1 DDX41 as an emerging DDR-associated protein.....	44
5.2 DDX41 facilitates DNA repair.....	45
5.3 DDX41 resolves R-loops.....	46
5.4 DDX41: a potential co-transcriptional suppressor of transcription replication conflicts.....	47
5.5 Prospective molecular pathogenesis of R525H mutation in MDS and AML patients.....	47
6. CONCLUSIONS AND FUTURE WORK.....	49
6.1 Conclusions.....	49

6.2. Future work	50
7. REFERENCE LIST.....	51

LIST OF FIGURES

Figure 1. Structures of helicase superfamilies.....	1
Figure 2. Motifs and their potential functions in superfamily 2 DEAD-box RNA helicases.....	3
Figure 3. Structure and sequence of human DDX41.....	4
Figure 4. Summary of DDX41 germline and somatic mutations reported in MDS and AML patients.....	6
Figure 5. DNA damage and repair mechanisms.....	8
Figure 6. The framework of DNA damage response signaling pathways.....	9
Figure 7. The structure of an R-loop.....	10
Figure 8. Pervasive R-loop formation leads to lethal transcription-replication collisions, a major threat to genome stability.....	11
Figure 9. Schematic representation of relationship between double-strand breaks and R-loops.....	12
Figure 10. IR induced foci formation of DDX41 in HT1080 cells.....	23
Figure 11. IR induced foci formation of DDX41-GFP in HT1080 cells.....	24
Figure 12. Altered protein levels are linked to DNA damage response in DDX41 KO cells.....	25
Figure 13. Prolonged DNA damage in DDX41 KO cells.....	26
Figure 14. Impaired DNA damage response in DDX41-KO HT1080 cells.....	27
Figure 15. Impaired DNA damage response in DDX41-KO HeLa cells.....	28
Figure 16. DDX41 colocalizes with γ H2AX upon DNA insult.....	30-31
Figure 17. Alkaline comet assay shows higher and prolonged DNA migration in DDX41-KO HT1080 cells.....	32
Figure 18. Dot blot assays shows accumulation of RNA:DNA hybrids in DDX41-KO cells.....	34
Figure 19. Purification of DDX41 proteins.....	35

Figure 20. Helicase and annealing activities of DDX41-WT and R525H mutant proteins on RNA:DNA hybrids.....	36
Figure 21. DDX41 colocalizes with R-loops upon DNA insult.....	38
Figure 22. DDX41 colocalizes with RNA Pol II upon bleomycin treatment.....	39
Figure 23. Overexpression of the R525H mutant results in prolonged DNA damage and diminishes repair in DDX41-KO HT1080 cells.....	41
Figure 24. Overexpression of the R525H mutant perturbs RNA:DNA hybrids in DDX41-KO HT1080 cells.....	42
Figure 25. Overexpression of R525H mutant perturbs RNA:DNA hybrids in wildtype HT1080 cells.....	43
Figure 26. A proposed model of the role of DDX41 in co-transcriptional resolution of R-loops to maintain genome stability.....	48

LIST OF ABBREVIATIONS

AML	Acute myeloid leukemia
ANKRD26	Ankyrin repeat domain-containing protein 26
ATG2B	Autophagy-related 2B
ATM	Ataxia-telangiectasia mutated
ATP	Adenosine triphosphate
ATR	ATM- and Rad3-related
ATRIP	ATR interacting protein
BER	Base-excision repair
Bm	Bleomycin
bp	Basepair
BRCC	BRCA1-BRCA2-containing complex
BSA	Bovine serum albumin
BTK	Bruton's tyrosine kinase
CEBPA	CCAAT enhancer-binding protein α
CHK2	Checkpoint kinase 2
CPD	Cyclobutane pyrimidine dimer
CSB	Cockayne syndrome B
cis-Pt	Cisplatin
CML	Chronic myeloid leukemia
DEAD	Asp-Glu-Ala-Asp
DDR	DNA damage response
DDX41	DEAD-box helicase 41
DMEM	Dulbecco's modified eagle's medium
DNA-PK	DNA-dependent protein kinase
DSB	Double-strand break
EDTA	Ethylenediamine tetraacetic acid
ETV6	ETS variant transcription factor 6
FANCA	Fanconi anemia complementation group A
FANCD2	Fanconi anemia complementation group D2

FBS	Fetal bovine serum
FL	Follicular lymphoma
FPLC	Fast protein liquid chromatography
γ H2AX	γ -H2A histone family member X
GATA2	GATA-binding factor 2
Gy	Gray
HLTF	Helicase-like transcription factor
HR	Homologous recombination
HRP	Horseradish peroxidase
HSPC	Hematopoietic stem and progenitor cell
HU	Hydroxyurea
IPTG	Isopropyl β -D-1-thiogalactopyranoside
IR	Infrared radiation
KO	Knockout
LB	Lysogeny broth
MDS	Myelodysplastic syndrome
MMC	Mitomycin C
MMR	Mismatch repair
NE	No enzyme
NER	Nucleotide-excision repair
NHEJ	Non-homologous end-joining
Ni-NTA	Nickel-nitrilotriacetic acid
NLS	Nuclear localization signal
NTP	Nucleotide triphosphate
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate-buffered saline
53BP1	P53 binding protein 1
pBRCA1	Phospho-BReast CAncer gene 1
PCNA	Proliferating cell nuclear antigen
Pi	Phosphate ion or inorganic phosphate
PMSF	Phenylmethanesulfonylfluoride

(6-4)PP	6–4 photoproduct
PVDF	Polyvinylidene difluoride
RAD52	Radiation sensitive 52
RF	Replication fork
RNAPII	RNA polymerase II
RNF8	Ring finger protein 8
RPA	Replication protein A
RUNX1	Runt-related transcription factor 1
SDS	Sodium dodecyl sulphate
SF	Superfamily
SRP72	Signal recognition particle 72
STING	Stimulator of interferon genes
TA-HRR	Transcription-associated homologous recombination repair
TBE	Tris/Borate/EDTA
TEMED	N,N,N',N'- tetramethylethylenediamine
TERC	Telomerase RNA component
TERT	Telomerase reverse transcriptase
TOPBP1	Topoisomerase 2 binding protein 1
TRCs	Transcription and replication collisions/conflicts
XPD	Xeroderma pigmentosum complementation group D
XPG	Xeroderma pigmentosum complementation group G
WT	Wild type

INTRODUCTION

1.1 Helicases

Helicases are the ubiquitous enzymes that transduce the chemical energy generated by hydrolysis of nucleoside triphosphate (NTP) into an oligonucleotide strand separation and displacement activity (Caruthers and McKay, 2002). Depending on substrates, helicases can be classified as DNA or RNA helicases, and based on polarity, are classified as 5'→3' or 3'→5' helicases (Patel and Picha, 2000). Helicases are engaged in almost all aspects of nucleic acid metabolism, including processes such as repair, replication, transcription, recombination,

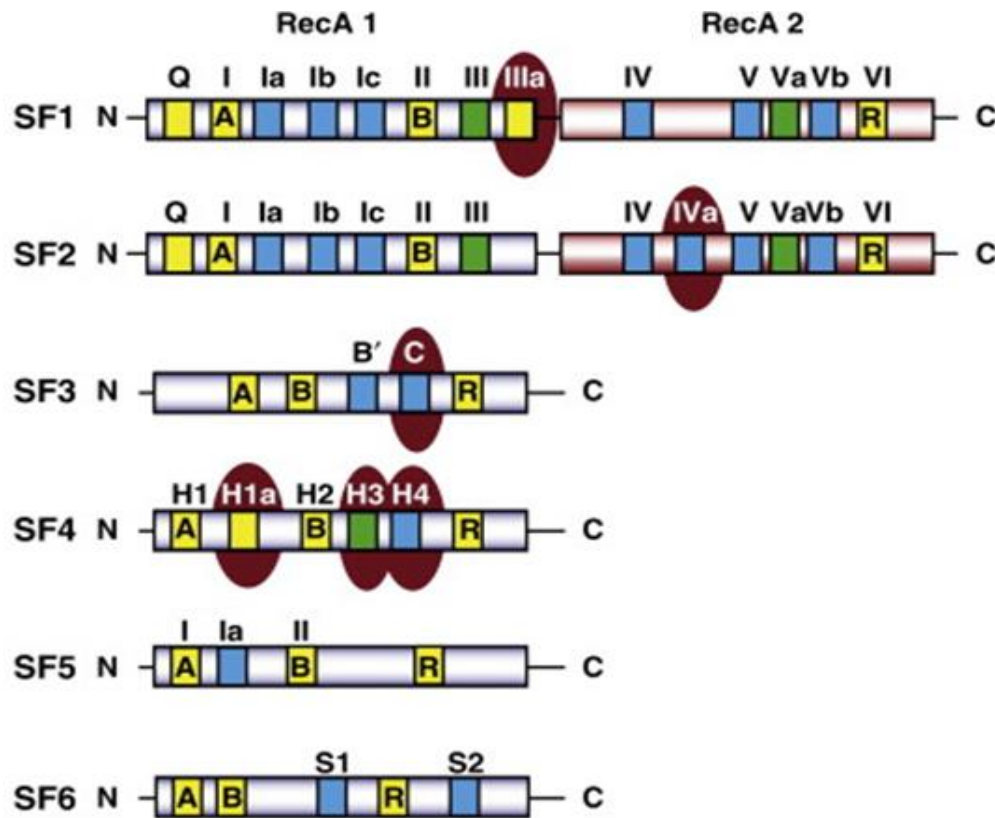


Figure 1. Structures of helicase superfamilies. An illustration of the core helicase domain. A blue cylinder represents the N-terminal RecA domain (RecA1), and the C-terminal RecA domain (RecA2) is presented as a red cylinder. Conserved amino acid motifs are coloured depending on the helicase function. Motifs in yellow are implicated in NTP binding/hydrolysis, motifs in green are correlated with translocation, and motifs in blue interact with the nucleic acid. Distinct motifs from all the illustrated superfamilies are highlighted with a red oval. Motif I (Walker A), motif II (Walker B), and arginine finger (R) are conserved throughout all helicase superfamilies. Taken from (Jackson et al., 2014).

chromosome segregation, and telomere maintenance (Bernstein et al., 2010; Brosh Jr and Bohr, 2007; Dillingham, 2011; Jankowsky, 2011). Helicases have been classified into six major groups, i.e., superfamily 1 (SF1)-SF6, based upon conserved amino acid sequence motifs (Singleton et al., 2007) (**Figure 1**). Among them, only the SF1 and SF2 comprise nine short conserved amino acid sequences, named Q, I, Ia, Ib II, III, IV, V, and VI, respectively, and SF2 is the largest group among these six superfamilies (Byrd and Raney, 2012; Gorbalenya and Koonin, 1993). These motifs are generally grouped in a section of 200–700 amino acids known as the helicase core domain. The helicase family is also called the DEAD-box (or DEAH or DEXH) protein family, because of the sequence of motif II (Asp-Glu-Ala-Asp, DEAD) (Lohman et al., 2008). Furthermore, X-ray crystallographic studies demonstrate that the conserved helicase motifs are linked with the tertiary structure of helicase proteins, suggesting that they form a significant functional domain, coordinating ATP binding and hydrolysis to nucleic acids unwinding (Lohman et al., 2008; Singleton et al., 2007).

1.2 DEAD-box helicases

The DEAD-box helicase is the largest subfamily under superfamily 2 RNA helicases. The DEAD motif, together with a motif I, Q motif, and motif VI, is required for ATP binding and hydrolysis (Pause et al., 1994; Pause and Sonenberg, 1992). Moreover, the DEAD motif, along with motif I, is important for magnesium binding. Motifs Ia and Ib, IV, and V are involved in intramolecular arrangement and RNA interaction, necessary for remodelling activity of the RNA helicase. Furthermore, motif III is essential for coupling of ATP hydrolysis to remodeling events like unwinding (Parsyan et al., 2011) (**Figure 2**).

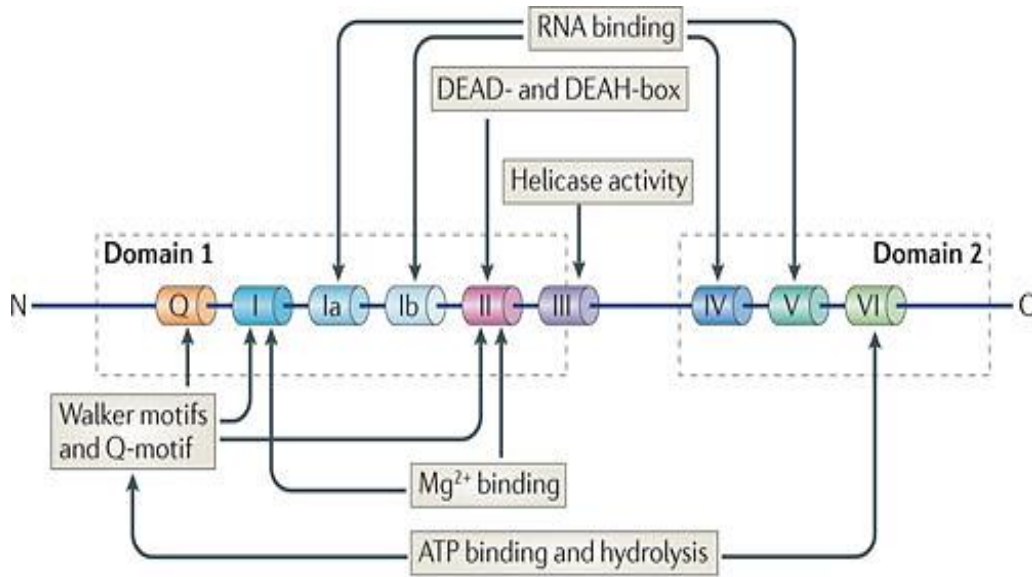


Figure 2. Motifs and their potential functions in DEAD-box RNA helicases. Taken from (Parsyan et al., 2011).

Several members of the DEAD-box helicases have drawn the high interest of researchers, as each of the individual enzymes from these subfamilies is engaged in several processes of RNA metabolism. For example, Ded1 and p72 engage in transcription and pre-mRNA splicing (Auboeuf et al., 2002; Liu et al., 2002), Drs1p, Dhr1p, and Dhr2p in ribosomes biogenesis (Venema and Tollervy, 1999), Dbp5 in nuclear export (Gatfield et al., 2001), Ded1p in translation initiation (Chuang et al., 1997), SKI2 and LGP2 in RNA degradation (Anderson and Parker, 1998; Py et al., 1996), and mHel61p in organelle gene expression (Missel et al., 1997). Moreover, these helicases are often linked to several diseases, such as malignancy of tumor cells and a few human male infertility syndromes (Fuller-Pace, 2013; Lasko, 2013). Further, DEAD-box helicases illustrate a large family of proteins that possess a non-processive and local dissociation activity on dsRNA substrates (Linder, 2006). This activity is primarily utilized in various RNA metabolic processes by eukaryotes. In these processes, the DEAD-box helicases likely act as valuable check-point proteins in the RNA metabolism machinery (Linder, 2006).

1.3 DDX41 helicase

DEAD-box helicase 41 (DDX41) is a member of the SF2 family (**Figure 3**). DDX41 is expressed in a large range of species, from *drosophila* to plants and mammals. DDX41 is composed of two domains: the N-terminal helicase core domain and the C-terminal Zinc-finger like domain. A nuclear localization signal (NLS) motif is present before the helicase domain. These domains further comprise different functional motifs that function in binding and hydrolysis of ATP, recognition of nucleic acids, and unwinding DNA/RNA unwinding (Linder and Fuller-Pace, 2013). Mass spectroscopy studies have revealed DDX41 is engaged in the second step of splicing where the 5' and 3' exon ligation and the release of the

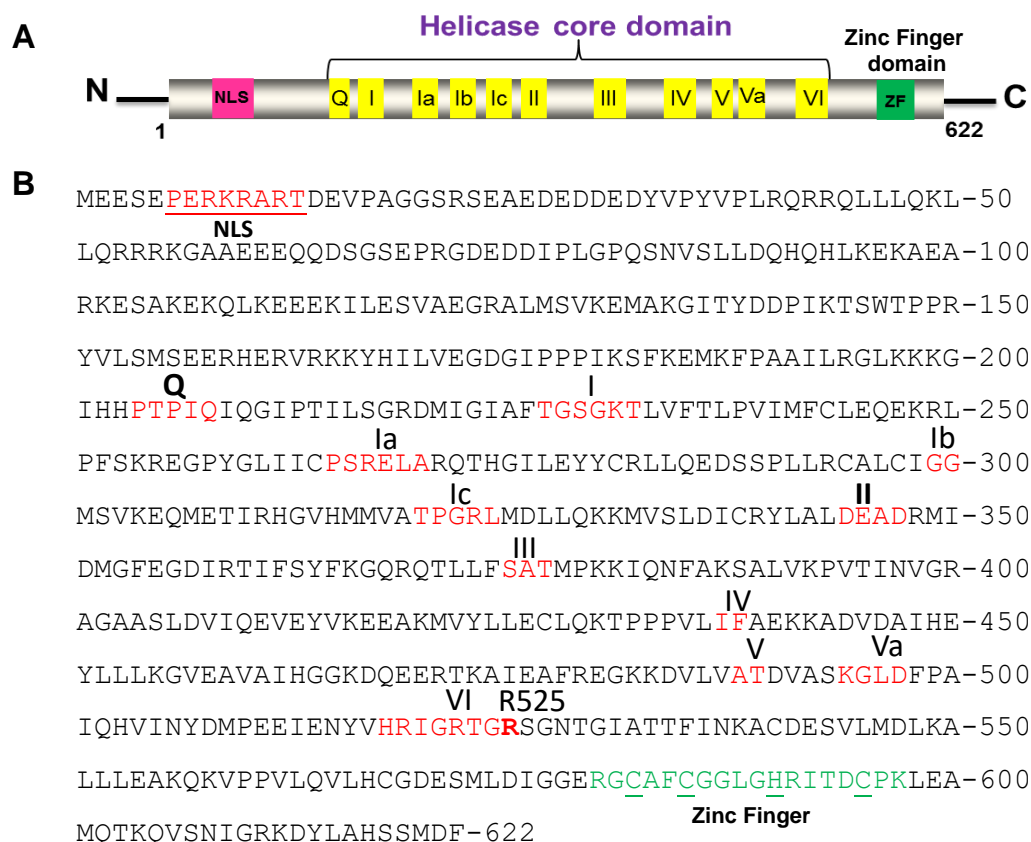


Figure 3. Structure and sequence of human DDX41 protein. (A) Schematic of DDX41 protein representing the nuclear localization sequence (NLS, in red), conserved helicase motifs (yellow), and zinc finger domain (ZF, green). (B) The sequence of human DDX41 protein.

intronic lariat occur (Agafonov et al., 2011; Bessonov et al., 2008; Jurica et al., 2002). A siRNA screening found that DDX41 functions as an intracellular DNA sensor in myeloid dendritic cells (Zhang et al., 2011b). In contrast to other Pathogen recognition receptors (PRRs, a group of germ line-encoded receptors), DDX41 can also recognize bacterial secondary messengers such as cyclic-di-GMP (c-di-GMP) or cyclic-di-AMP (c-di-AMP) to activate Type I interferon immune response (Parvatiyar et al., 2012). It has been shown that Bruton's tyrosine kinase (BTK) phosphorylates DDX41 at Tyr414 and activates it, thereby, allowing for recognition of dsDNA after viral or bacterial infection (Lee et al., 2015b). Following ligand recognition, DDX41 binds STING (Stimulator of Interferon Genes), and STING subsequently phosphorylates the TANK-binding protein and interferon regulatory factor 3 (IRF3) and finally induces the production of type I interferon (Ishikawa and Barber, 2008; Jiang et al., 2017; Omura et al., 2016; Tanaka and Chen, 2012).

Besides its role in innate immunity, both germline and acquired somatic mutations in DDX41 have been associated with myelodysplastic syndrome (MDS) and acute myeloid leukemia (AML) (Cardoso et al., 2016; Ding et al., 2012; Fenwarth et al., 2021). The most frequently identified somatic mutation is c.1574G>A, p.R525H (67%), which is mainly confined to the highly conserved region of DDX41 (Quesada et al., 2019; Lewinsohn et al., 2016; Li et al., 2016b; Polprasert et al., 2015). Moreover, mutations in DDX41 are also linked with acute erythroid leukemia (AEL) (Iacobucci et al., 2019). DDX41 mutations also lead to failure of tumor suppressor function such as aberrant pre-mRNA splicing, RNA processing, and cell cycle arrest (Kadono et al., 2016; Lewinsohn et al., 2016; Polprasert et al., 2015; Qin et al., 2021). Another study suggested DDX41 mutants may be involved in activation of the DNA damage response (DDR), resulting in cell cycle arrest (Peters et al., 2017). Although DDX41 is well studied as an essential sensor in the field of the innate immune response, its potential role in DDR remains unexplored.

1.4 MDS, AML and DDX41 mutations

MDS and AML are hematologic diseases prevalent in adults 65 years of age or older (Almeida and Ramos, 2016; Ria et al., 2009). MDS is characterized by unsuccessful hematopoiesis and peripheral blood cytopenia (a disorder in which the number of blood cells are lower than normal), whereas in AML, bone marrow is characterized by the spread of immature myeloid

cells. Progression of MDS increases the risk of its evolution into AML (Jiang et al., 2009). To date, eleven genes, namely ANKRD26, ACD, ATG2B, CEBPA, DDX41, ETV6, GATA2, RUNX1, SRP72, TERC, and TERT that are indicative of germline heterozygous mutations. These mutations have been identified to be associated with familial MDS and AML (Duployez et al., 2016; Lee et al., 2015a; Tawana et al., 2018; Tawana and Fitzgibbon, 2016), of which DDX41 is the latest (Jahn et al., 2018; Maciejewski et al., 2017). High-throughput sequencing of samples from MDS and AML patients has revealed several MDS and AML-related mutations (Bejar et al., 2011; Haferlach et al., 2008; Mossner et al., 2016; Papaemmanuil et al., 2013); among those, the DDX41-R525H mutation was one of the most recurrent ones (**Figure 4**). This R525H mutation has been revealed to have a reduction in nucleotide coordination (Li et al., 2016b), inhibitory effect on cell cycle progression by repressing E2F (E2F Transcription Factor) activity, defect in pre-rRNA processing, and diminished protein synthesis as a consequence of ribosomal stress (Kadono et al., 2016).

Generally, MDS and AML progression results from elevated levels of oncogenes, deactivation of tumor suppressor genes, or a combination of both, which then prompts replication stress and genome

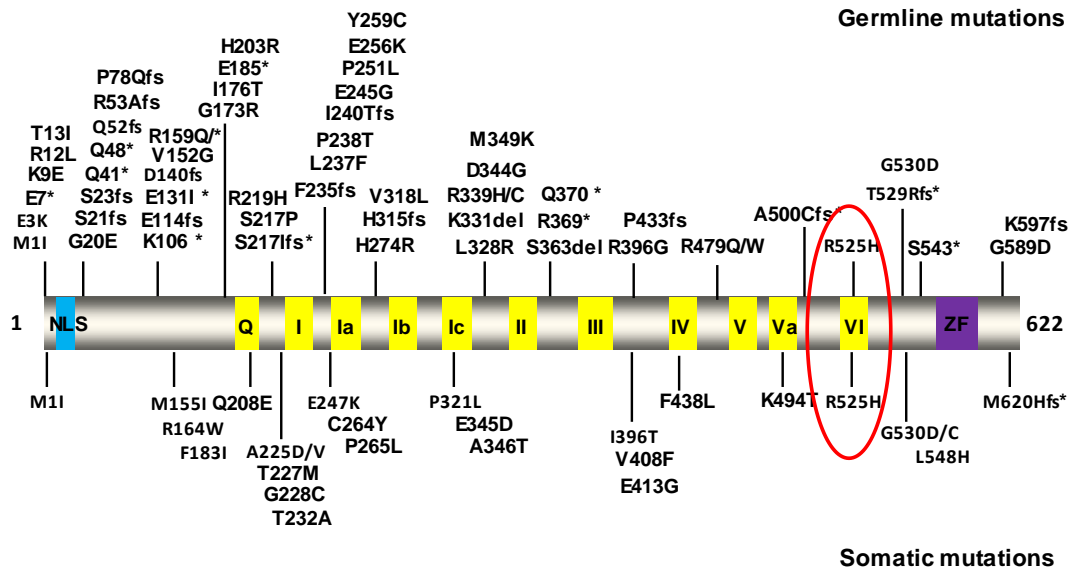


Figure 4. Summary of DDX41 germline and somatic mutations reported in MDS and AML patients. Nuclear localization signal (NLS, in blue), helicase motifs (yellow), and zinc finger (ZF, purple) domains are indicated. The R525H mutant is highlighted.

instability (Haferlach et al., 2008; Halazonetis et al., 2008). Several studies focus on alterations in DNA damage response (DDR) associated with MDS and AML in which DNA double-strand breaks (DSBs) are critical DNA damage and genome instability markers (Boehrer et al., 2009; Cavelier et al., 2009; Horibe et al., 2007; Kefala et al., 2013). Several DNA damage response genes, such as ATM (Grosjean-Raillard et al., 2009), XPD and XRCC1 (Joshi et al., 2016), BRCC (Meyer et al., 2020), and HLTf (Takaoka et al., 2019), have been associated with MDS/AML. Some DEAD-box proteins, such as DDX1 (Li et al., 2008; Zhang et al., 2011a) and DDX3 (Sun et al., 2013; Szappanos et al., 2018), are also involved in DDR. DNA damage measured by γ H2AX and 53BP1 (DSB markers) increases through MDS to AML; in addition, there is diminished DDR in MDS and AML cell lines and samples of bone marrow (Popp et al., 2017), suggesting that dysregulated DDR and DSB might be a potential root of MDS and AML pathogenesis.

1.5 DNA damage and repair

Cells are constantly subjected to DNA damage, generating thousands of DNA lesions per cell each day (Ciccia and Elledge, 2010). A variety of lesions are found in DNA from three leading causes. The first leading cause is environmental agents like ultraviolet radiation, ionizing radiation, and several genotoxic chemicals (Ding et al., 2012). The second leading cause is cellular metabolic products, like reactive oxygen species originated from oxidative respiration and lipid peroxidation products (Hoeijmakers, 2001). Thirdly, some chemical bonds in DNA disintegrate under a few physiological conditions, like hydrolysis of nucleotide residues, which leaves non-instructive abasic sites; for example, deamination of cytosine is converted to the miscoding uracil (Chatterjee and Walker, 2017). DNA lesions may sometimes also occur due to replication errors caused by tautomeric shifts (the "imino" or "enol" form of nucleotides), eventually resulting in base-pair mismatching such as an A with a G instead of a T (Tubbs and Nussenzweig, 2017) (**Figure 5**).

DNA damage response pathways and DNA repair proteins are utilized by the cell to repair, remove, or tolerate a variety of DNA lesions, to limit the genomic instability (Hoeijmakers, 2001). As there are different types of DNA lesions, a single repair process cannot mitigate all the classes of DNA damage. At least five main DDR pathways function in mammals: nucleotide-excision repair (NER) removes helix-distorting lesions, base-excision

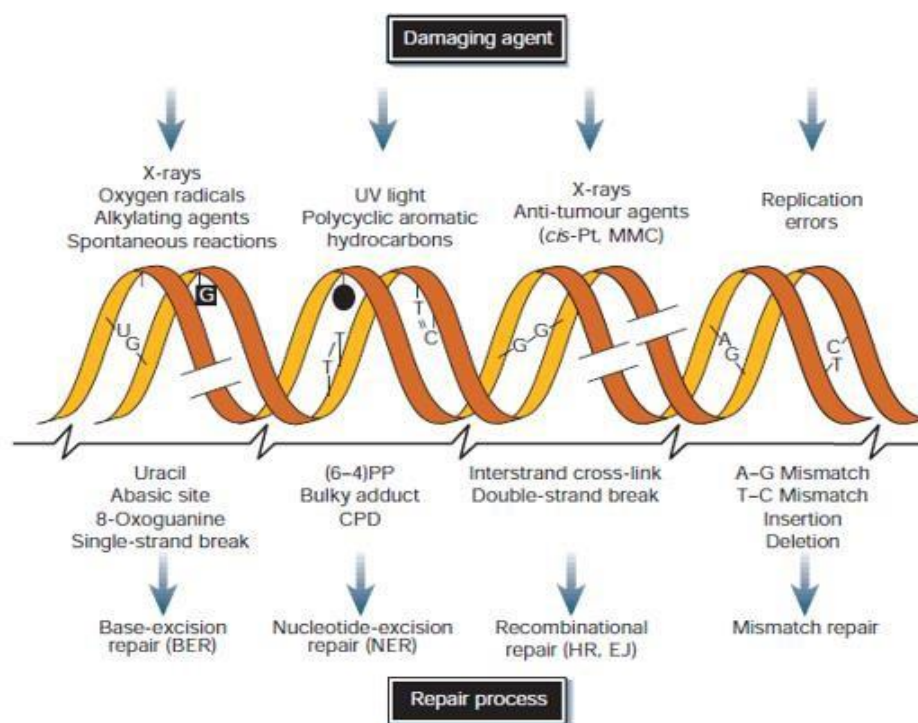


Figure 5. DNA damage and repair mechanisms. At the top, various DNA damaging agents, in the middle, examples of DNA lesions caused by these agents, and at the bottom, associated DNA repair mechanisms responsible for the removal of the lesions. cis-Pt, cisplatin; MMC, mitomycin C; (6-4)PP, 6-4 photoproduct; CPD, cyclobutane pyrimidine dimer. Taken from (Hoeijmakers, 2001).

repair (BER) reverses oxidative base modifications, homologous recombination (HR) and non-homologous end-joining (NHEJ) are involved in repair and removal of DSBs through distinct mechanisms (Lindahl and Wood, 1999), and mismatch repair (MMR) helps in restoring errors that occurred during replication (Jiricny, 2006) (**Figure 5**). DDR uses signal sensors, transducers, and effectors similar to classic signal transduction pathways (**Figure 6**) (Zhou and Elledge, 2000).

Unlike signal transduction pathways activated by ligands of receptor kinases, the DDR signaling pathway is activated by anomalous DNA structures (Maréchal and Zou, 2013). The DNA damage signaling pathways are mainly accelerated by the transducers such as ATM (ataxia-telangiectasia mutated), ATR (ATM- and Rad3-Related), and DNA-PK (DNA-dependent protein kinase). Sensors such as the MRN complex (MRE11, RAD50, and NBS1) detect DSBs, then recruit and activate ATM at the lesion site (Ryan et al., 2016). Activated ATM phosphorylates several effector proteins, such as CHK2 (Checkpoint kinase 2), H2AX, 53BP1, and p53. Phosphorylation of H2AX by ATM and ubiquitination of H2AX

by activators, namely RNF8 (Ring Finger Protein 8) and RNF168 E3 ubiquitin ligases, lead to the recruitment of DNA repair proteins like BRCA1/pBRCA1 (Turnell and Grand, 2012) (**Figure 6**). Alternatively, ATR is primarily activated at ssDNA regions first coated by the sensor RPA (Replication protein A) (Ball et al., 2005).

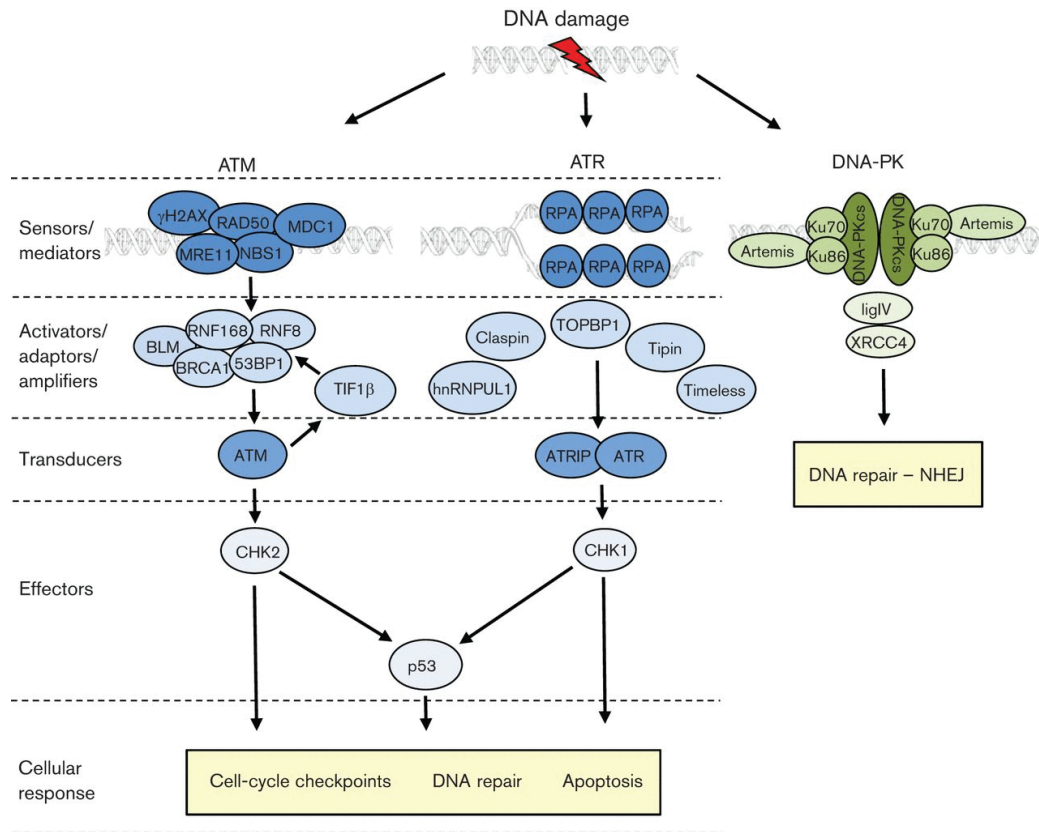


Figure 6. The framework of DNA damage response signaling pathways. Various protein complexes acting distinctly as either sensors, transducers, activators, or effectors in ATM, ATR and DNA-PK pathways. The arrows represent the flow of the individual pathways. Taken from (Turnell and Grand, 2012).

Further, ATR is recruited to RPA-coated ssDNA through its interacting protein ATRIP (ATR interacting protein) (Namiki and Zou, 2006). The recruitment of activator TOPBP1 (Topoisomerase 2 Binding Protein 1) to the ssDNA region is responsible for the ATR activation that later phosphorylates effector proteins like CHK1 (Ball et al., 2005). These effectors ultimately result in various cellular responses such as DNA repair, cell cycle arrest, and apoptosis. Unlike the ATM and ATR kinases, DNA-PK is responsible for regulating the NHEJ repair pathway (Turnell and Grand, 2012). Since DDX41 has not been implicated in the DDR,

it remains unknown whether DDX41 acts as a DDR sensor, transducer, or effector. Moreover, the mechanism by which DDX41 may potentially produce DDR is an unanswered question.

1.6 R-loops

R-loops originate during transcription, when nascent RNA exits RNA polymerase, and pairs with its complementary DNA template, resulting in RNA:DNA hybrid structure and displaced single-stranded DNA (Chakraborty et al., 2018a) (**Figure 7**). R-loops are found in many organisms where they perform various cellular processes, involving regulation of chromosome segregation, replication of bacterial plasmids and mitochondrial genomes (Kabeche et al., 2018), and accomplishing immunoglobulin class-switch recombination (Aguilera and García-Muse, 2012). In mammalian cells, R-loops take up almost 5% of the genome (Ginno et al.,

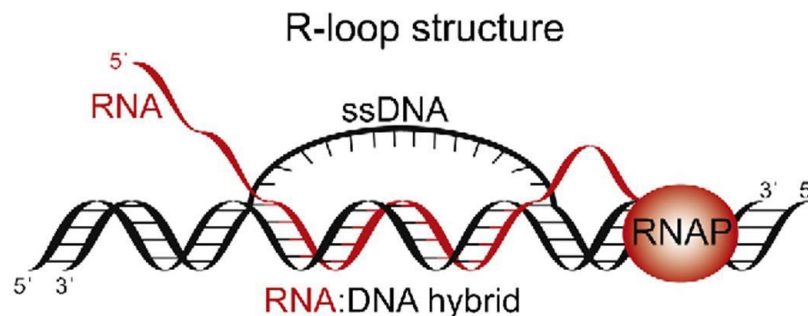


Figure 7. The structure of an R-loop. ssDNA, single-stranded DNA; RNAP, RNA Polymerase. Taken from (Hamperl and Cimprich, 2014).

2012). They are abundant at promoter and terminator regions of polyA-dependent genes, which implies the potential role of R-loops in the regulation of gene expression (Ginno et al., 2012). Further, R-loops are also found in tRNA and rDNA genes, implying their formation through transcription involving different RNA Polymerases (Santos-Pereira and Aguilera, 2015).

Despite the various regulatory role of R-loops, these can also cause a major threat to genomic stability if they persists, i.e., generating single-stranded, which can adversely lead to stalling or collapsing of replication forks (Basu et al., 2011; Hamperl and Cimprich, 2014; Santos-Pereira and Aguilera, 2015) and fatal double-strand DNA breaks by the endonucleases XPF and XPG, involved in nucleotide excision repair (NER) pathway. (Sollier et al., 2014). The R-loops formed during transcription are also deciphered to promote the aggravation of DNA damage by escalating transcription and replication collisions (TRCs) (Sanchez et al., 2020). At times but

not exclusively, the machinery of transcription and replication processes results in a collision and consequently leads to increased DNA breaks as both the processes utilize the same template of DNA (Sanchez et al., 2020) (**Figure 8**). Persistent R-loop structures and the TRCs lead to genomic instability that prompts various cellular fates (Lang et al., 2017) and diseases in humans, including neurodegenerative diseases (Chen et al., 2018; Haeusler et al., 2014; Loomis et al., 2014) and MDS (Chen et al., 2018).

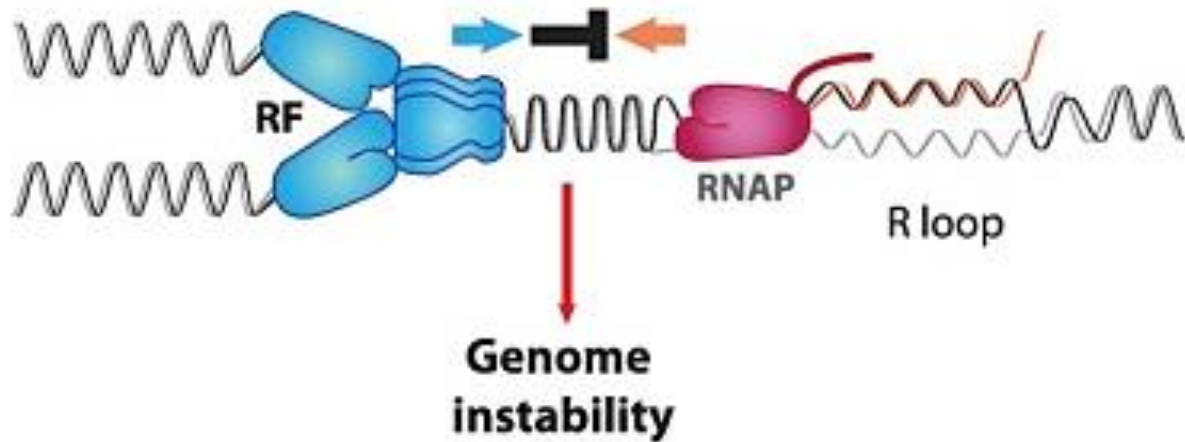


Figure 8. Pervasive R-loop formation leads to lethal transcription- replication collisions, a major threat to genome stability. RF, Replication fork; RNAP, RNA polymerase II. Taken from (García-Rubio et al., 2018).

Besides transcription, R-loops are widely associated with DSBs to both induce and inhibit DNA repair, depending on whether the R-loops formation is transient or persistent. A study shows that in many instances, induction of DSBs leads to R-loops, which then recruit transcription-associated homologous recombination repair (TA-HRR) factors such as RAD52 and XPG (Kato et al., 2019). Such mechanisms are crucial for protecting genome integrity, particularly in transcriptionally active regions. Alternatively, R-loops also promote DSB formation by activating the ATM associated DDR pathway (Sollier and Cimprich, 2015). Upon the extreme R-loop removal, the efficiency of two central DSB repair pathways, HR and NHEJ, was reduced (Lu et al., 2018). This study led researchers to affirm the importance of R-loop structures in promoting DSB repair (Ohle et al., 2016; Yasuhara et al., 2018). Conversely, as mentioned above, persistent and excessive R-loops may challenge the DDR pathway by forming pathological DSB and hindering DNA repair in different ways. In one research study, overexpression of RNase H1, an enzyme that removes R-loops by explicitly degrading

RNA:DNA hybrid (Cerritelli and Crouch, 2009) as well as active ribosomes and anti backtracking mechanisms, suppressed DSB accumulation (Dutta et al., 2011), which suggests that extension of R-loops is an important cause of DSBs, and in the long run, genome instability (Aguilera et al., 2012). Another consideration is that DSBs formed due to various DNA damage sources also induce R-loops, which can trigger both DNA repair or damage depending on whether it is moderate and transient or excessive and persistent, respectively (**Figure 9**). In the former case where R-loops are resolved and DNA is repaired, the process begins with an increase in transcription induced at DSB sites (Ohle et al., 2016), leading to a rise in R-loop structures. These structures then recruit DNA repair factors like CSB and carry forward the

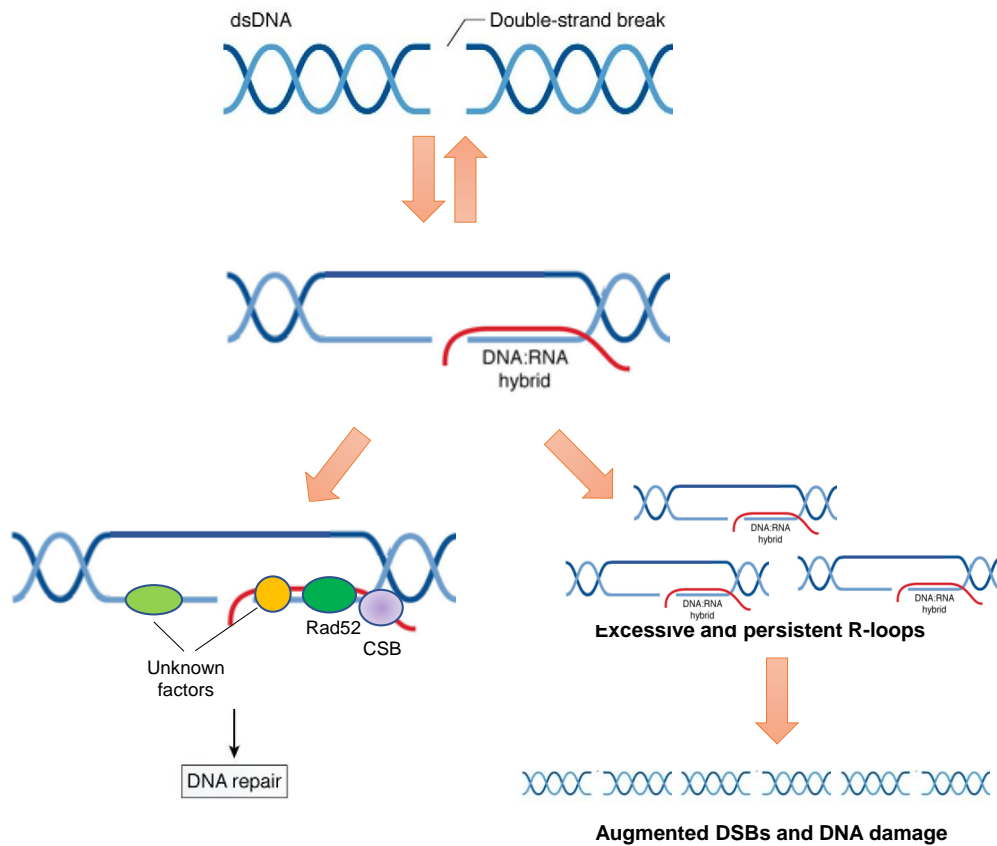


Figure 9. Schematic representation of the relationship between double-strand breaks and R-loops. R-loops can be formed as a consequence of DSBs and at situations R-loop formation leads to DSBs. In both situations, if excess of either R-loops or DSBs persists due to loss of repair proteins and factors, it leads to elevated DNA damage. CSB, Cockayne syndrome protein B. Modified from (Hegazy et al., 2020).

DDR pathway (Hegazy et al., 2020). On the other hand, R-loops accumulation and the inability

of repair factors to resolve R-loops results in DNA damage, intensification in DSB, and genome instability. Moreover, another research study has revealed that an increase in R-loop formation explains how the mechanism of mutations in splicing factors leads to MDS; in particular, mutations in splicing factors SRSF2 and U2AF35 augment R-loop formation at gene promoters (Chen et al., 2018).

Various proteins hold vital positions to resolve R-loop structures, which may otherwise lead to secondary pathological structures, mutagenesis, DNA damage-inducing TRCs. FANCA and FANCD2 (Fanconi anemia pathway factors) (Gaillard et al., 2015; Schwab et al., 2015) DNA/RNA helicases and various types of RNase H (Kim and Jinks-Robertson, 2012) are some of them. Interestingly, a recent study using zebrafish showed that depletion or mutation of DDX41 caused the accumulation of R-loops, which triggered the inflammatory signals and caused increased production of hematopoietic stem and progenitor cells (HSPCs) (Weinreb et al., 2021). Thus, DDX41 appears as a potential co-transcriptional DDR protein that might resolve R-loops and avoid TRCs to maintain genome stability and prevent the occurrence of MDS and AML pathogenesis.

2. HYPOTHESIS AND OBJECTIVES

2.1 Hypothesis

DDX41 is required for efficient DNA damage response and R-loop resolution to maintain genome stability.

2.2 Objectives

1. Determine the potential role of DDX41 in DNA damage response.
2. Determine the mechanisms of DDX41 in DNA repair response.
3. Determine co-localization of DDX41 with DNA double-strand break marker, transcription machinery, and R-loop in cells using confocal microscopy.
4. Examine the levels of RNA:DNA hybrid in DDX41 KO cells-reconstituted with DDX41-WT or DDX41-R525H gene.

3. Methods

3.1 Plasmid DNA

DDX41 (both full-length) was PCR amplified and cloned into the Hind III and BamH I sites of a pAcGFP1-N2 vector (Clontech). All point mutations were generated with a QuikChange Site-Directed Mutagenesis Kit (Agilent Technologies).

For CRISPR-knockout, four sgRNAs used targeting exons 2 and 3 of the human DDX41 gene were designed with <http://crispr.mit.edu>, the pair of DNA oligonucleotides were synthesized by IDT, then annealed and ligated into the Bbs I site of pSpCas9(BB)-2A-Puro (pX459) vector (#62988, Addgene), as described (Ran et al., 2013). The insert was confirmed with Age I and Bbs I digestion. LentiCRISPR v2 plasmids targeting the DDX41 gene were constructed using methods as described (Sanjana et al., 2014). All plasmids were verified by DNA sequencing.

3.2 Cell lines

HeLa and HT1080 cells (from ATCC) were grown in DMEM supplemented with 10% FBS with penicillin and streptomycin (100 U/mL each, Sigma-Aldrich). CRISPR DDX41 knockout HeLa and HT1080 cell lines were generated by Dr. Ravi Shankar Singh (Post-doc) in Dr. Yuliang Wu's lab (University of Saskatchewan); Next, lentivirus overexpression of DDX41-WT and DDX41-R525H gene in wildtype and DDX41-KO background cells have also been established by Dr. Ravi Shankar Singh; To overexpress both WT and R525H DDX41 in DDX41 KO or WT cell lines, *DDX41* gene was cloned into *Kpn* I and *Not* I sites of the entry vector pEN_TTmcs (25755, Addgene), and transferred to the destination vector pSLIK Hygro (25737, Addgene) by the Gateway system. The lentivirus particles were prepared as described above, and the expression of DDX41 was induced by doxycycline 20 ng/mL, Invitrogen

3.3 CRISPR knockout

For both HT1080 and HeLa cells transfection, 1×10^6 cells were seeded in a 6-well plate 24 h before transfection. Two μ g of DNA was transfected to each well using Lipofectamine 3000 (Invitrogen). After 18 h of transfection, the cells were transferred into a 10 cm dish containing

5 µg/mL puromycin (Sigma) and maintained in selection medium until clones were visible. The clones were picked individually and grown in 6-well plates until confluency, with the cells lysed using RIPA buffer (25 mM Tris, pH 7.4; 150 mM NaCl; 0.5% sodium deoxycholate; 1% NP-40) with 2 mM phenylmethylsulfonyl fluoride (PMSF) and protease inhibitor cocktail (Sigma). The gene KO was confirmed by Western blot analysis, followed by PCR amplification of the target segment, and cloning into TOPO-TA vector (Invitrogen) for sequencing.

3.4 Reagents

Table 1. List of reagents, catalog number and suppliers

Reagents, Cat No.	Suppliers	Address
Acrylamide, 0314	AMRESCO	North York, Ontario, Canada
N,N'-Methylenebisacrylamide, AC164790250hrp	Fisher Scientific	Madison, Wisconsin, USA
Adenosine Triphosphate (ATP), A1852	Sigma-Aldrich	Oakville, Ontario, Canada
Clarity Max Western ECL substrate, 1705062	Bio-Rad Laboratories, Inc.	Mississauga, Ontario, Canada
Agarose, 5510UB	Gibco	Burlington, Ontario, Canada
Ampicillin, A1593	Sigma-Aldrich	Oakville, Ontario, Canada
Ammonium persulfate (APS), A3678	Sigma-Aldrich	Oakville, Ontario, Canada
Boric acid, SLBC5554V	Sigma-Aldrich	Oakville, Ontario, Canada
Bovine serum albumin, A2058	Sigma-Aldrich	Oakville, Ontario, Canada
Bradford protein assay reagent, 500-0013	Bio-Rad	Hercules, California, USA
Bromophenol blue, B0126	Sigma-Aldrich	Oakville, Ontario, Canada
Chloramphenicol, AC227920250	Fisher Scientific	Madison, Wisconsin, USA
Dithiothreitol (DDT), 10197777001	Sigma-Aldrich	Oakville, Ontario, Canada
Dulbecco's Modified Eagle Medium (DMEM), SH30022.01	Sigma-Aldrich	Oakville, Ontario, Canada
EDTA, 17892	Fisher Scientific	Madison, Wisconsin, USA
Fetal Bovine Serum (FBS), F6178	Sigma-Aldrich	Oakville, Ontario, Canada
[γ - ³² P] ATP, BLU502A	Perkin Elmer	Boston, MA, USA
Glycine, BP381-5	Fisher Scientific	Madison, Wisconsin, USA
Glycerol, 123170	Fisher Scientific	Madison, Wisconsin, USA

Imidazole, O3196-500	Fisher Scientific	Madison, Wisconsin, USA
Isopropyl β -D-1-thiogalactopyranoside (IPTG), 15528019	Fisher Scientific	Madison, Wisconsin, USA
Kanamycin, 60615	Fisher Scientific	Madison, Wisconsin, USA
Lysogeny Broth (LB) with agar, L2897	Sigma-Aldrich	Oakville, Ontario, Canada
Methanol, 154246	Sigma-Aldrich	Oakville, Ontario, Canada
Magnesium chloride (MgCl ₂), M8266	Sigma-Aldrich	Oakville, Ontario, Canada
Nickel-NTA Affinity beads, 70666	Sigma-Aldrich	Oakville, Ontario, Canada
Non-fat dry milk, 170-6404	Sigma-Aldrich	Oakville, Ontario, Canada
N,N,N',N'-Tetramethylethylenediamine (TEMED), 87689	Bio-Rad	Hercules, California, USA
Penicillin-Streptomycin, 084M4778V	Sigma-Aldrich	Oakville, Ontario, Canada
Polyethyleneimine (PEI) MAX, 24765	Polysciences, Inc.	Warrington, Pennsylvania
Phenylmethylsulfonyl fluoride (PMSF), P7626	Sigma-Aldrich	Oakville, Ontario, Canada
Protein standard, 161-0374	Bio-Rad	Hercules, California, USA
Protease inhibitor, 05892791001	Roche	Mannheim, Germany
Polyvinylidene difluoride membrane (PVDF), 10600023	Bio-rad	Mississauga, Ontario, Canada
Sodium dodecyl sulfate (SDS), L3771	Sigma-Aldrich	Oakville, Ontario, Canada
Sodium chloride (NaCl), S671-10	Fisher Scientific	Madison, Wisconsin, USA
Thiamine hydrochloride, T4625	Sigma-Aldrich	Oakville, Ontario, Canada
Tris, 0826	AMRESCO	North York, Ontario, Canada
Tris (2-carboxyethyl) phosphine, C4706	Sigma-Aldrich	Oakville, Ontario, Canada
TritonTMX-100, X100	Sigma-Aldrich	Oakville, Ontario, Canada
Trypsin-EDTA, T4049	Sigma-Aldrich	Oakville, Ontario, Canada
Tryptone, TRP402.205	BioShop	Burlington, Ontario, Canada
TWEEN 20, BP337	Fisher Scientific	Madison, Wisconsin, USA

3.5 DNA damage treatments

An equal number of 2×10^6 WT and DDX41-KO HeLa and HT1080 cells were seeded in a 6-

well plate using a cell counter and incubated at 37°C, 5% CO₂ for 16 hours in 2 mL of DMEM medium. After 16 hours, DNA damage was induced by exposure to 5 Gy ionizing radiation (IR produced by X-Rad225 XL, Precision X-Ray), 10 mM hydroxyurea (HU, VWR cat # CAAAA10831-03) for 16 h, 50 J/m² of UV radiation, or 30 µg/mL of Bm (Bleomycin sulphate, Bioshop cat# BLE011.10) for 16 hours.

3.6 Antibodies

DDX41 mouse antibody (MABF1107) from Sigma-Aldrich was used for certain immunofluorescence imaging, and DDX41 rabbit antibody (D3F1Z) from New England Biolabs (NEB) was used for both immunofluorescence imaging and immunoblotting, which recognizes the N-terminal of DDX41 (98 amino acid of DDX41). γH2AX (Ser139, 20E3), phosphor- BRCA1 (Ser1524, 9009), and 53BP1 antibodies (E7N5D) were from NEB. S9.6 antibody was from Kerafast (ENH001). Double-stranded DNA antibody (ab27156) and RNA polymerase II antibody (ab817) were from Abcam. β-actin-HRP conjugated mouse antibody (sc-47778 HRP) was from Santa-Cruz. Goat anti-rabbit IgG HRP-conjugated (7074) and goat anti-mouse IgG HRP-conjugated (7076) secondary antibodies were from NEB. Immunofluorescence secondary antibodies, IgG (H+L) highly cross-adsorbed goat anti-rabbit Alexa Fluor 488 (A11008) and goat anti-mouse Alexa Fluor 594 (A-11032), were from Invitrogen.

3.7 Confocal microscopy

Cells seeded onto coverslips in a 6-well plate were washed twice with PBS and fixed with 100% methanol at -20°C for 30 min, then washed with PBS and blocked with blocking buffer (1% BSA in PBST, PBS+0.1% Tween 20) at room temperature for 1 h. The immunostaining was performed by incubating cells with primary antibodies (1:100) in blocking buffer overnight at 4°C. After washing with 1 mL of PBS (thrice), the cells were incubated with secondary antibodies, Alexa Fluor 488 goat anti-rabbit IgG (1:1000, Invitrogen), and Alexa Fluor 594 goat anti-mouse IgG (1:1000, Invitrogen) in blocking buffer for 1 h at room temperature. Cells were then washed with 1 mL of PBS (thrice) and mounted with Prolong Diamond antifade reagent containing DAPI (Invitrogen) and cured at room temperature in the

dark for 16-20 hours. Immunofluorescence was performed on a Zeiss LSM 700META inverted Axiovert 200 M laser scan microscope with a Plan-Apochromat 63×/1.4 oil DIC objective. Images were captured with a CCD camera and analyzed using LSM Browser software ZEN (Zeiss).

For GFP tagged proteins, cells were directly fixed with methanol, mounted with Prolong Diamond antifade reagent containing DAPI and then observed under an LSM 700 microscope.

3.8 Western blotting

Depending on the size of the target protein, proteins were separated on 6, 8, 10, 12, or 15% SDS- polyacrylamide gel using Tris-glycine buffer (with 10 % SDS). After electrophoresis, the proteins were transferred to a PVDF membrane (Bio-Rad) using Wet/Tank blotting system in the transfer buffer (25 mM Tris, 192 mM glycine, pH 8.3, and 20% methanol) at 100 V for 3 h at 4°C. The membrane was incubated with blocking buffer (5% BSA in PBST) for 1 h at room temperature. The membrane was then incubated in the primary antibodies of interest (1:1000, prepared with 1% BSA in PBST) at 4°C, for 16 hours. After washing with 5 mL of PBST five times (5 min each), the membrane was incubated with HRP conjugated secondary antibody (1:5000, NEB) with 1% BSA in 5 mL of PBST for 1 h at room temperature; then washed three times with 5 mL of PBST (5 min each). The membrane was then treated with Clarity Western ECL Substrate (Bio-Rad) and visualized using a ChemiDoc MP Imaging System (Bio-Rad).

3.9 Alkaline comet assays

An alkaline comet assay was performed using a CometAssay kit (R&D systems, cat# 4250-050-K) according to the manufacturer's instructions. Briefly, 2×10^4 cells/mL were seeded in 6-well plates using a cell counter and incubated at 37°C, 5% CO₂ for 16 hours in DMEM medium, and treated the next day with bleomycin (30 µg/mL, 16 hours) and collected at different time points, post-treatment. Cells were washed with 1 mL of PBS and centrifuged at 4000 rpm for 5 mins, and the pellet was mixed with 50 µL of PBS and 500 µL of molten low-melting-point agarose at 37°C. For the PBS and agarose mixture, 50 µL was applied onto microscope slides that were

precoated with agarose and incubated at 4°C for 30 min in the dark (to avoid any further DNA damage in the cells). To unwind the supercoiled DNA, cells were lysed using the CometAssay Lysis Solution (from the CometAssay kit) for 1 h at room temperature in the dark and then exposed to an alkaline solution (pH >13) for 20 min at room temperature in the dark. Next, the electrophoresis was performed using Horizontal Electrophoresis System from Bio-Rad at 21 volts for 30 min, where the DNA-strand breaks present in the cells migrate in the direction of the anode. Finally, the slides were washed with water, fixed with 70% ethanol for 15 min, and then stained with GelRed (Biotium) for 30 min at room temperature in the dark. DNA was visualized under a confocal microscope (Zeiss LSM 700) at a 20× magnification. The open-source software ImageJ (<https://imagej.nih.gov/ij>) was used to measure the tail length.

3.10 Dot blot assays

On day 1, cells (5×10^6) were collected and lysed in 1 mL lysis buffer, composed of 500 μ L 2× lysis buffer (0.5% SDS, 40 mM Tris-HCl pH 7.5, 150 mM NaCl, and 5 mM EDTA), 500 μ L 2x TE (100 mM Tris-HCl pH 8, 10 mM EDTA pH 8), and 10 μ L of proteinase K (20 mg/ml), at 37°C for 16 hours. Genomic DNA was then extracted using DNeasy Blood & Tissue Kit (Cat # 69506, Qiagen) and samples of extracted DNA (~50 μ g each) fragmented using a restriction enzyme cocktail of 1 μ L HindIII (20 U/ μ L), 1 μ L EcoRI (20 U/ μ L), 2 μ L BsrGI (10 U/ μ L), 1 μ L XbaI (20 U/ μ L), and 4 μ L SspI (5 U/ μ L, all from NEB) in NEB Buffer 2 (in 300 μ L total) at 37°C for overnight. On day 3, the mixture was treated with 2 U of RNaseIII (NEB) and 4 U of RNase T1 (Fisher) for 2 h. The digested fragments were divided into two parts: one was treated with 4 U of RNase H for 2 h, and another remained untreated. The fragmented DNA samples were repurified using the DNeasy Blood & Tissue Kit and dissolved in 100 μ L of 5 mM Tris-HCl pH 8.5. The extracted DNA was quantified using a Nanodrop (Thermo Fisher). Each sample was diluted to 250 ng in 50 μ L of TE and was spotted on a nylon membrane and crosslinked by UV treatment. The membrane was blocked with 5% milk in 10 mL of PBST and incubated with S9.6 antibody (1: 2000 dilution in 3% BSA of PBST) for 16 hours at 4°C. The membrane was then washed five times with 10 mL of PBST and incubated with secondary antibody (1: 1000 dilution in 3% BSA of PBST) for 1 h, and later the membrane was treated with ECL reagent (Bio-Rad) and visualized using a ChemiDoc MP Imaging System (Bio-Rad). The same membrane was stripped with 10 mL of stripping buffer for 15-20 mins

(15 mg/mL glycine, 1 mg/mL SDS, and 1% Tween 20, pH 2.2), re-blocked with 10 mL of 5% BSA in PBST for 1 hour, and then incubated with dsDNA antibody (1:2000 dilution in 3% BSA of PBST) for 1 h and then washed and next incubated with secondary antibody (1: 1000 dilution in 3% BSA of PBST) for 1 h. After three times of washing with 10 mL of PBST for 5 minutes each, the nylon membrane was visualized by the ChemiDoc MP Imaging System.

3.11 Recombinant proteins

The plasmid containing pDEST17-DDX41-WT or pDEST17-DDX41-R525H mutant was transformed into *E. coli* Rosetta 2 cells (DE3, EMD Millipore). The cells were grown at 37°C in 1 L of LB medium containing 100 µg/mL of ampicillin and 34 µg/mL of chloramphenicol until an OD₆₀₀ was reached 0.6. Cells were then induced with 0.3 mM IPTG for 16 hours at 15°C. The cells were harvested by centrifugation at 5000 g for 10 min at 4 °C. The periplasmic material was removed from the cells as described (Magnusdottir et al., 2009). Briefly, the cells were suspended in 5 mL/g (cell mass) of hypertonic buffer (50 mM HEPES, pH 7.4, 20% sucrose, and 1 mM EDTA) on ice for 30 min and centrifuged at 8000 g for 10 min at 4 °C. The cells were re-suspended in 5 mL/g of cell mass with hypotonic solution (5 mM MgSO₄) and incubated for 10 min on ice. Cells were then pelleted by centrifugation at 4000 g for 10 min at 4 °C and stored at -80 °C until use. The cells were lysed by sonication in buffer A (50mM HEPES-NaOH, pH 7.0, 300 mM NaCl, 20 mM imidazole, 10% glycerol, 2 mM MgCl₂, and 5 mM β-mercaptoethanol) with a final concentration of 1 mM PMSF and 10µL of 100x protease inhibitor (RocheApplied Science) at 4 °C, with 10 short bursts of 10 s at intervals of 10 min. The cell debris and inclusion bodies were removed by centrifugation at 45,000 g for 30 min at 4 °C. Recombinant His-tagged DDX41 proteins were subjected to a two-step purification involving Nickel Affinity beads (Sigma-Aldrich) and a Sephacryl S-300 HR 16/60 gel filtration column (GE Healthcare). The supernatant was applied to the Ni-NTA beads equilibrated with buffer A, washed with 10 column volumes (CVs) of buffer B (50 mM HEPES-NaOH, pH 7.0, 150 mM NaCl, 10% glycerol, 2 mM MgCl₂, and 5 mM β-mercaptoethanol) containing 20 mM imidazole, and proteins were eluted with 5 CVs of buffer C (50 mM HEPES-NaOH, pH 7.0, 300 mM NaCl, 10% glycerol, 2 mM MgCl₂, 5 mM β-mercaptoethanol, and 500 mM imidazole). The composition of protein fractions were evaluated by SDS-PAGE, and the fractions with high recombinant protein content were pooled and subjected to size-exclusion

chromatography on a Sephacryl S-300 HR 16/60 column (GE Healthcare) that was equilibrated and eluted with buffer D (50 mM HEPES-NaOH, pH 7.0, 150 mM NaCl, 10% glycerol, 2 mM MgCl₂, and 1 mM DTT). The fractions were collected at a flow rate of 0.5 mL/min with the same buffer. The protein composition of each peak was assessed by SDS-PAGE, and the fractions of interest were pooled and concentrated.

All proteins were snap-frozen in liquid nitrogen and stored at -80°C. Protein concentration was determined by the Bradford method (Bradford, 1976) using bovine serum albumin (BSA) as the standard.

3.12 RNA and DNA substrates

PAGE-purified oligonucleotides were purchased from Integrated DNA Technologies (IDT):

RNA 30 mer: 5'-GAGCTACCAGCTACCCCGTATGTCAGAGAG-3' and

DNA 30 mer comp+15T:

5'-CTCTCTGACATACGGGGTAGCTGGTAGCTCTTTTTTTTTTTTTTTTTTTTTTTTTTTT-3'

A single oligonucleotide for each substrate was 5'-end-labeled with [γ -³²P] ATP (Perkin Elmer) at 37°C using T4 polynucleotide kinase (NEB) for 1 h. Unbound radionucleotides were removed by using a G25 chromatography column (GE Healthcare). For the RNA:DNA hybrid substrate, an [γ -³²P]ATP-labeled oligonucleotide was annealed to a 2.5-fold excess of the unlabeled complementary strands in annealing buffer (10 mM Tris-HCl, pH 7.5, 50 mM NaCl) by heating at 95°C for 5 min and then cooling slowly to room temperature. The annealed RNA:DNA substrates were stored at 4°C until used.

3.13 Helicase assays

Helicase assay reaction mixtures (20 μ L) comprised 40 mM Tris (pH 8.0), 0.5 mM MgCl₂, 15 mM NaCl, 0.01% Nonidet P-40, 0.1 mM DTT, 1 mg/mL BSA, an equimolar mixture of 2 mM ATP and MgCl₂, 0.5 nM of RNA:DNA substrate, and the indicated concentrations of DDX41 protein. Helicase reactions were initiated by adding DDX41 protein and then incubated at 37 °C for 15 min. Reactions were quenched with the addition of 20 μ L of 2 \times Stop buffer (17.5

mM EDTA, 0.3% SDS, 12.5% glycerol, 0.02% bromphenol blue, and 0.02% xylene cyanol). A 10-fold excess of unlabeled oligonucleotide (cold oligo that has the same sequence as the [γ - ^{32}P]ATP labelled strand) was included in the quench to prevent reannealing. The products of the helicase reactions were resolved on nondenaturing 15% (19:1 acrylamide: bisacrylamide) polyacrylamide gels. ^{32}P -radiolabeled DNA or RNA species in polyacrylamide gels were visualized using Phosphor-Imager Typhoon FLA 7000 (GE Healthcare).

3.14 Strand annealing assays

A 20ul reaction was carried out with 0.5 nM of RNA:DNA hybrid substrate. The substrate was first denatured at 100°C for 5 min and then incubated with 0 to 3 μM of DDX41 protein at 37°C for 15 min with or without 2 mM ATP in the helicase assay buffer (see above). After incubation, the reaction was stopped by addition of 20 μL of 2 \times Stop buffer (17.5 mM EDTA, 0.3% SDS, 12.5% glycerol, 0.02% bromophenol blue, and 0.02% xylene cyanol). The mixture was resolved on 15% native PAGE gel for 2 h at 180 V. The resolved radiolabeled species were visualized using the Phosphor-Imager Typhoon FLA 7000.

4. RESULTS

4.1 DDX41 forms foci upon IR exposure

To investigate the potential role of DDX41 in DNA damage response, I examined the ability of both endogenous (**Figure 10**) and exogenously expressed (**Figure 11**) (tagged with GFP) DDX41, to form foci in response to IR (ionizing radiation, 2 Gy) in HT1080 cells. In this work I largely address biological DNA damage and evaluated by measuring the foci yields. These subnuclear foci are formed upon various DNA damage related response factors, which accumulate explicitly at damaged sites, are spotted as distinct spots in nuclei (Rothkamm et al., 2015). IR directly affects DNA structure by inducing DNA breaks, particularly DSBs (Cannan and Pederson, 2016). DDX41 foci formation appeared after IR treatment, and the number of foci per cell increased over 4 h, i.e., an average of 5 at 0.5 h to 35 at 4 h of IR treatment (**Figure 10B and D**). In this experiment, IgG (**Figure 10A**) and γH2AX (**Figure 10C**) were used as negative and positive controls, respectively. As expected, no foci were formed in the IgG panel throughout the time before and after DNA damage, (**Figure 10A and D**) whereas γH2AX foci peaked 2 h after IR

treatment (**Figure 10A and D**). Next, I performed confocal imaging of exogenous DDX41 over 4 h after IR treatment to check any IR-induced foci formed. Here, Bloom helicase, which is a well-known DNA repair protein (Patel et al., 2017), was used as a positive control (**Figure 11D**). GFP vector (**Figure 11B**) and DDX41-K9A (**Figure 11C**) mutant (only localizes in the cytoplasm) acted as negative controls. I found that there was a 2-fold increase in the number of DDX41 foci from 0.5 h to 4 h (**Figure 11 A and E**). There was no foci formation observed in cells harbouring the GFP vector and DDX41-K9A. To elaborate, GFP vector was found both in cytoplasm and nucleus, without any specific localization and foci formation. Additionally, DDX41-K9A was predominantly localized in the cytoplasm and did not result in any IR-induced foci. Furthermore, Bloom helicase showed increased in foci over time, at 0 h the foci were minimal (i.e., 2 at an average), whereas it gradually increased to 6 times (relative to 0 h) by 4 h of IR treatment. Hence, these results suggest a possible role of DDX41 in DDR.

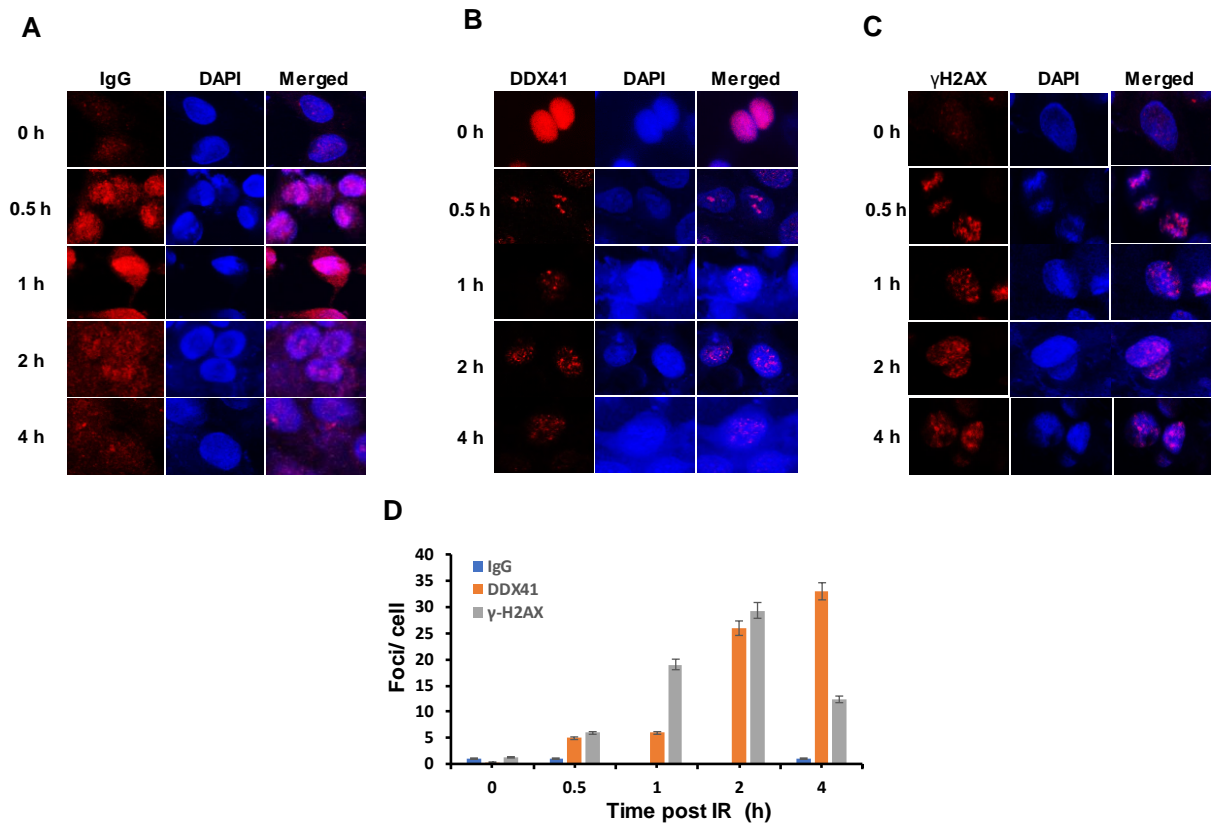


Figure 10. IR induced foci formation of DDX41 in HT1080 cells. Immunofluorescence of (A) IgG (negative control), (B) DDX41, and (C) γ H2AX (positive control), over time course of 0, 0.5, 1, 2, and 4 h after exposure to ionizing radiation in HT1080 cells. Nuclear DNA was counterstained with DAPI. (D) Quantitative analysis of foci from 100 cells shown in panels A-C in three biological replicates \pm SD.

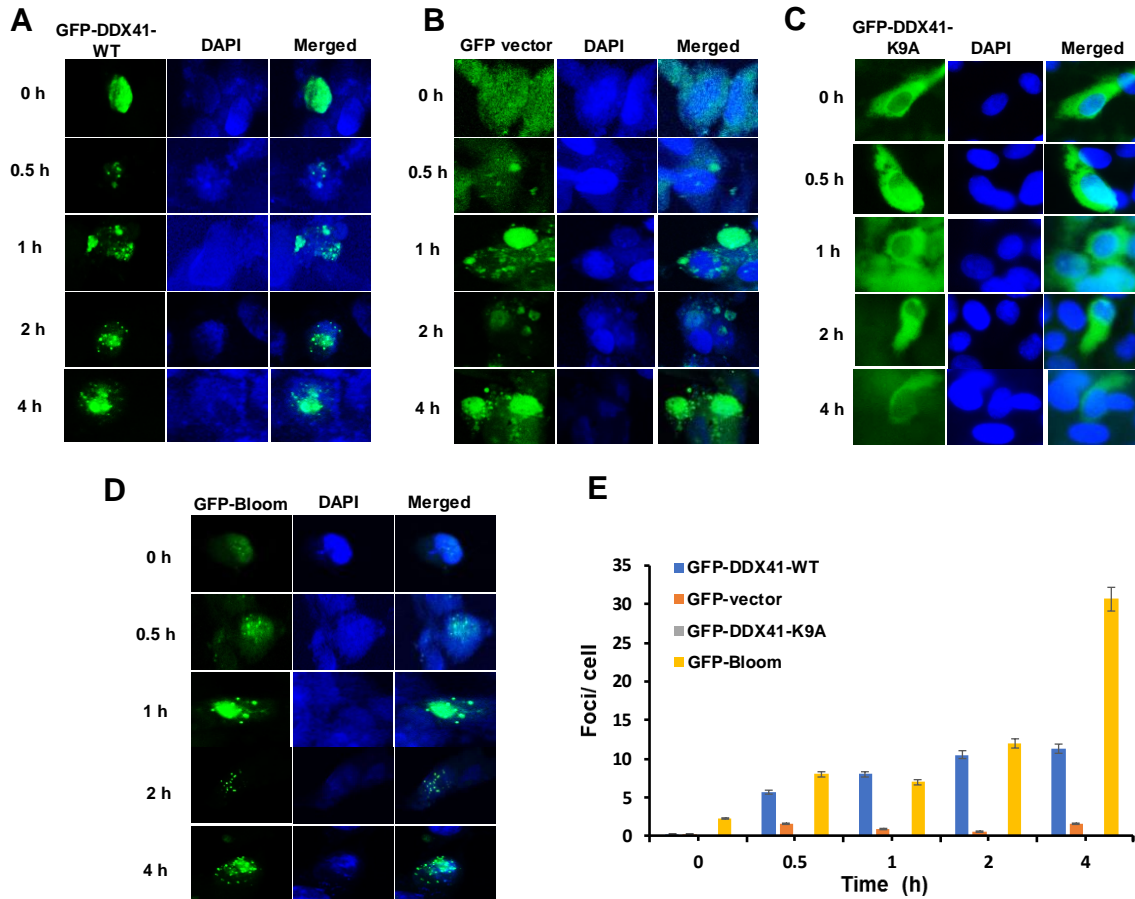


Figure 11. IR induced foci formation of DDX41-GFP in HT1080 cells. Immunofluorescence of (A) GFP-DDX41-WT, (B) GFP vector (negative control), (C) GFP-DDX41-K9A (mutant, negative control), (D) GFP-Bloom (positive control), over time course of 0, 0.5, 1, 2, and 4 h in HT1080 cells after exposure to 2 Gy of ionizing radiation. Nuclear DNA was counterstained with DAPI. (E) Quantitative analysis of foci from 100 cells shown in panels A-D in three biological replicates \pm SD.

4.2 DDX41 is involved in the DNA damage response pathway

To confirm the role of DDX41 in DDR and DNA repair, I performed immunoblotting to visualize some DDR related proteins (γ H2AX, 53BP1, p-CHK2, p-BRCA1, p-ATM, and p-P53) in both wildtype (WT) and DDX41-KO cell lines before and after treatment with IR (**Figure 12A**) or HU (**Figure 12B**) DNA damage for a time course of 24 h. Besides IR, I also included hydroxyurea (HU), which is used to model DNA damage by depleting the cells of dNTPs and resulting in stalled replication forks (Gottmann et al., 2010). I found that the levels of DNA damage proteins, γ H2AX and 53BP1, was increased from 0.5 h to 4 h in WT cells, whereas in DDX41-KO cell line

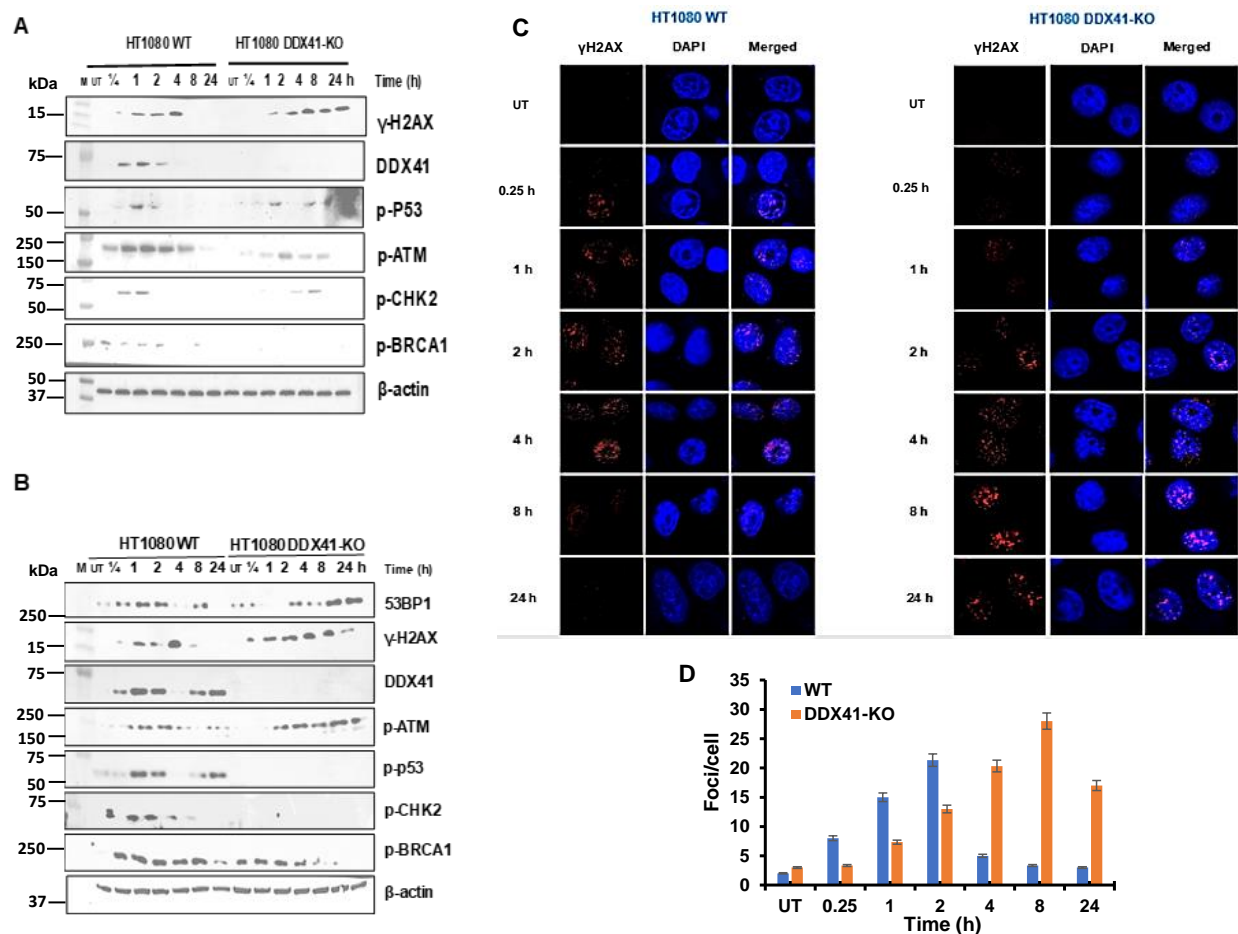


Figure 12. Altered protein levels are linked to DNA damage response in DDX41 KO. (A and B) Immunoblot analysis of DDR protein levels in WT and DDX41 KO HT1080 cells after IR (A, 5 Gy) and HU (B, 10 mM) treatment. β actin serves as a loading control. (C) Confocal images of γH2AX foci in WT and DDX41-KO cells after 5 Gy IR treatment, and collected at different time points i.e., at 15 min (1/4, 0.25 h), 1, 2, 4, 8, 24 h. (D) Quantification data of foci from 100 cells shown in panel C in three biological replicates ±SD. M, marker; UT, untreated; WT, wildtype; DDX41-KO, DDX41-knockout.

the γH2AX and 53BP1 level was comparatively higher and prolonged up to 24 h (Figure 12A and B). Additionally, the phosphorylation of DNA repair signaling protein, BRCA1 to pBRCA1 (Ser1524, DNA damage-induced phosphorylation site on BRCA1) was delayed and diminished in DDX41-KO cells. To elaborate, the damaged DNA repair process indicated by phosphorylation of BRCA1 initiated within 0.5 h after the DNA damage in WT cells, but the phospho-levels of BRCA1 were weak and appeared later around 8 h in the KO cell lines. Hence, this data bolstered my hypothesis of DDX41 functioning in the DNA damage response pathway.

I further verified these results by confocal imaging. I found that IR-induced γ -H2AX protein foci were observed up to 24 h in DDX41-KO cells, compared to 4 h in WT cells (**Figure 12C and D**), indicating that delayed or impaired DSB repair maybe related to DDX41 loss. To investigate the time at which DNA damage is resolved, I increased the investigation time to 72 h. I observed that the DNA damage indicated by γ -H2AX and 53BP1 levels was not resolved until 24 h after treatment in KO cells. In contrast, the DNA damage dissipated after 4 h in WT cells, suggesting that damaged DNA was repaired earlier in the presence of DDX41 (**Figure 13**). These data suggested DDX41 is required for DNA damage repair, and loss of DDX41 results in impaired and delayed DNA damage responses.

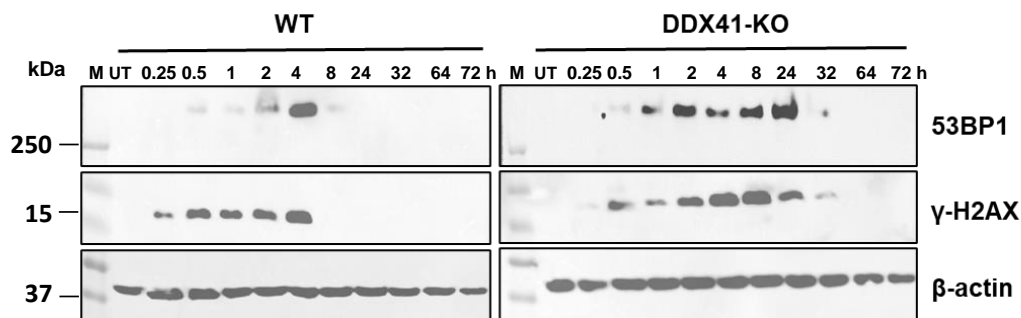


Figure 13. Prolonged DNA damage in DDX41 KO cells. Western blots showing protein levels of 53BP1 and γ -H2AX in WT and DDX41-KO HT1080 cells at different time points, i.e., at 0.25, 0.5, 1, 2, 4, 8, 24, 32, 64, and 72 h, after IR treatment. M, marker; UT, untreated; WT, wildtype; DDX41-KO, DDX41-knockout.

4.3 Increased DNA damage in DDX41-knockout cells

To verify the results above, I treated cells with additional DNA damage treatments, including IR, HU, and UV. UV radiation causes two classes of DNA lesions: cyclobutane pyrimidine dimers and 6-4 photoproducts (Kemp and Sancar, 2012). Immunoblots of WT and DDX41^{-/-} HT1080 (**Figure 14**) and HeLa (**Figure 15**) cell lines after IR (**Figure 14A and 15A**), UV (**Figure 14B and 15B**), or HU (**Figure 14C and 15C**) were performed to examine the DDR-related protein levels up to 72 h post-treatment. In agreement with the results described above, I found γ H2AX and 53BP1 levels were increased and prolonged in both HT1080 and HeLa DDX41-KO cells compared to the wild-type parental lines. Levels of

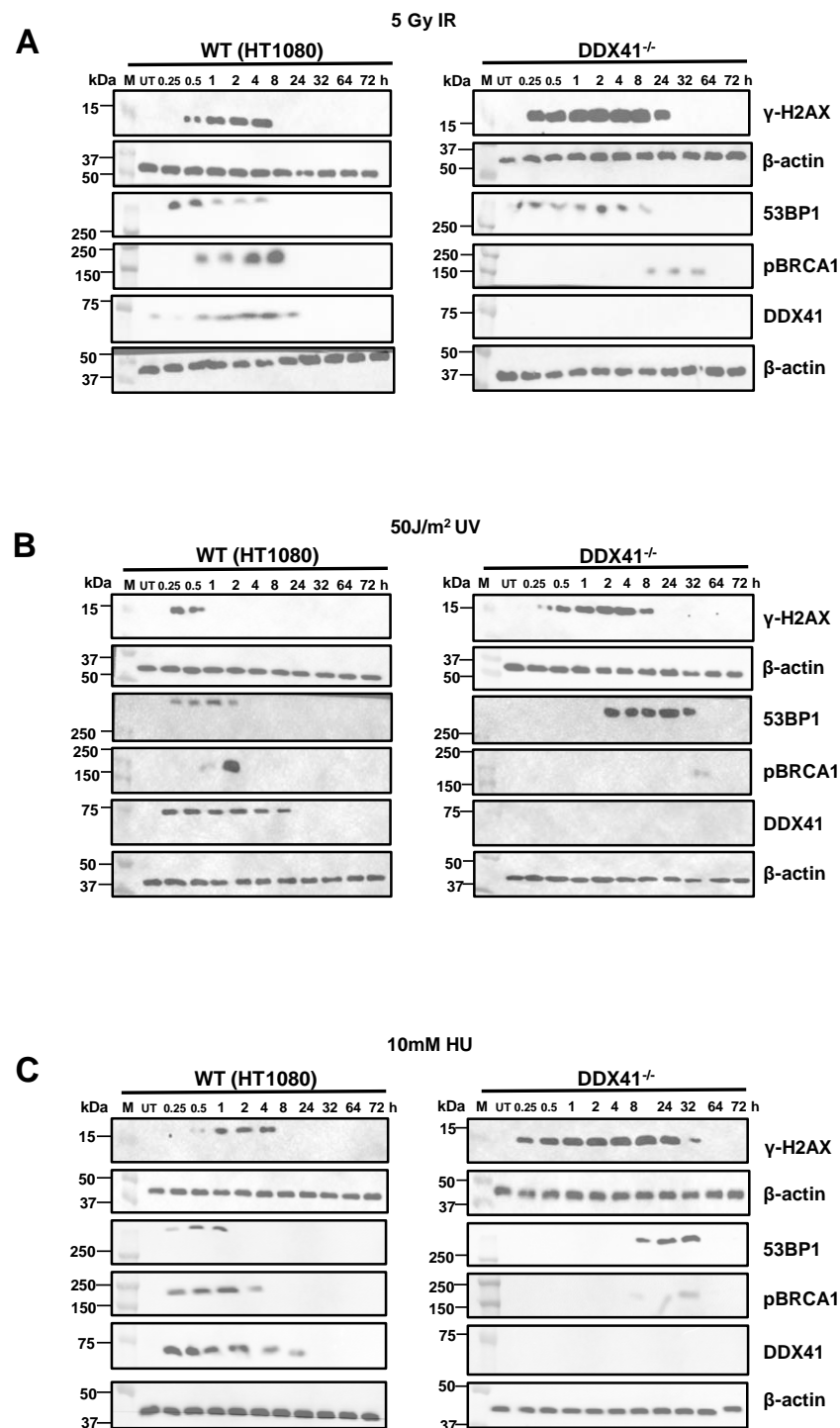


Figure 14. Impaired DNA damage response in DDX41-knockout HT1080 cells. Western blot assays of DDR-related proteins in WT and DDX41^{-/-} HT1080 cell lines after (A) IR (5 Gy), (B) UV (50J/m²), and (C) HU (10 mM for overnight) treatment at various time points (UT, 0.25, 0.5, 1, 2, 4, 8, 24, 32, 64, and 72 h) in three biological replicates. β-actin serves as a loading control.

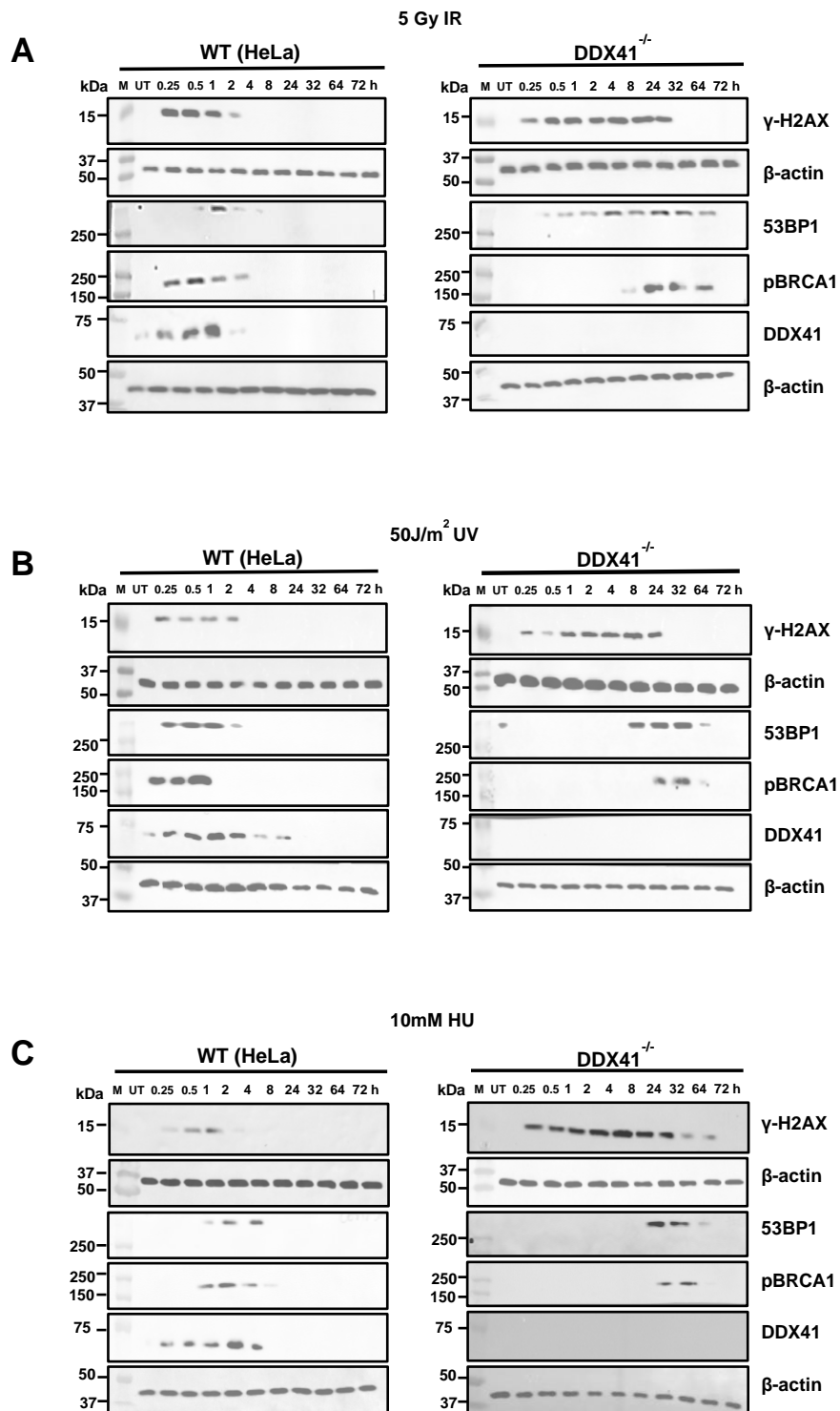
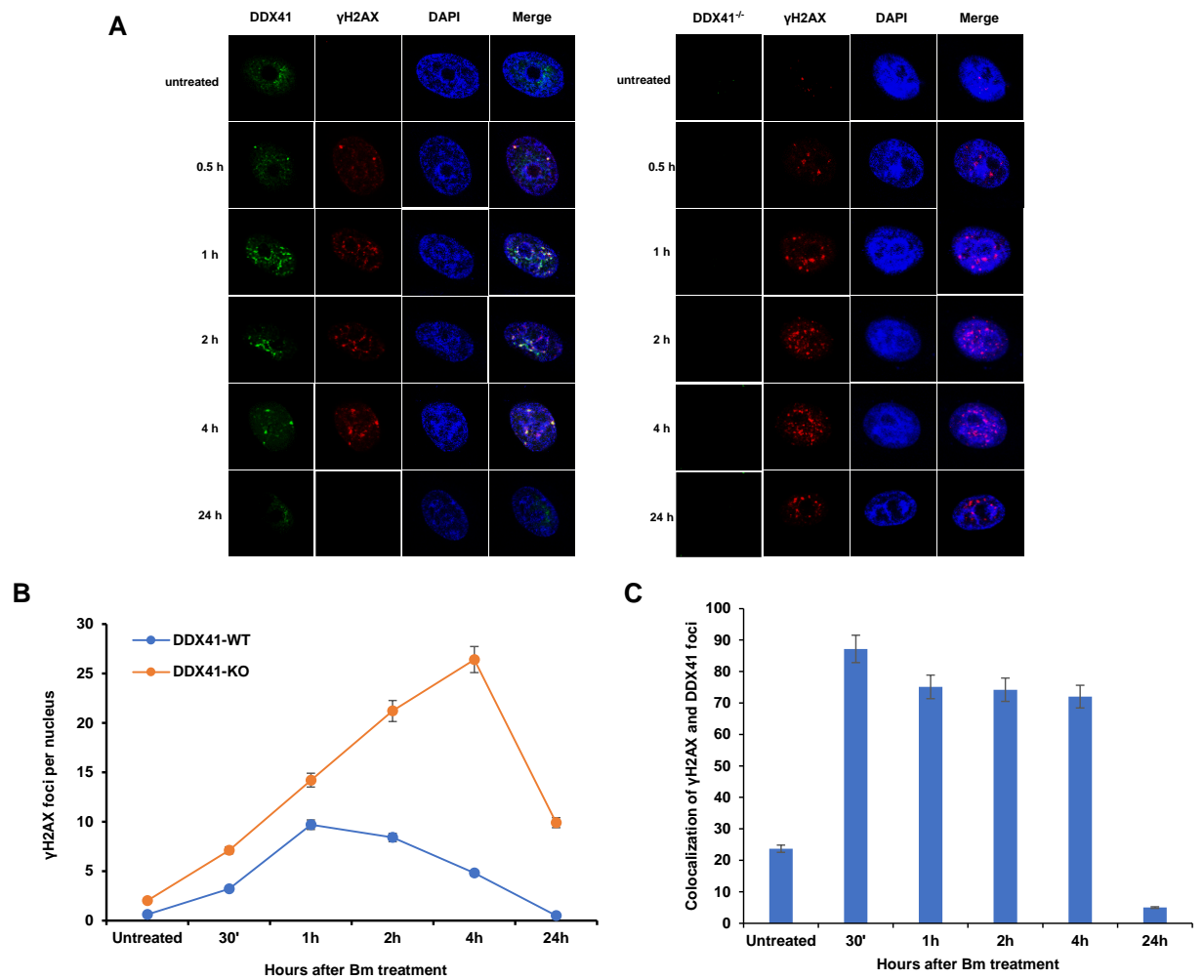


Figure 15. Impaired DNA damage response in DDX41-knockout HeLa cells. Western blot assays of DDR-related proteins in WT and DDX41^{-/-} HeLa cell lines after (A) IR (5 Gy), (B) UV (50J/m²), and (C) HU (10 mM for overnight) treatment at various time points (UT, 0.25, 0.5, 1, 2, 4, 8, 24, 32, 64, and 72 h) in three biological replicates. β-actin serves as a loading control.

53BP1 was majorly mirrored by the levels of pBRCA1 for a period of 1-4 h in WT cells compared to 24-64 h in KO cells, suggesting the significance of DDX41 in early DNA damage repair. Taken together, these results implied the crucial role of DDX41 in DDR and DSB repair.

4.4 DDX41 co-localizes with DNA double-strand break marker γ H2AX

To further investigate whether DDX41 is directly involved in DDR, I used confocal imaging to determine if DDX41 and γ H2AX co-localize following DNA damage. Because the IR machine used in above experiments was out of order for later experiments, I used bleomycin (Bm) to generate DSB. Bleomycin is a radiomimetic agent that causes DNA strand breaks, such as DSB (Chen et al., 2008). Intriguingly, I observed colocalization of DDX41 (green) and γ H2AX (red), ranging from 0.5 h to 2-4 h in both HT1080 (**Figure 16A and B**) and HeLa (**Figure 16D and E**) cell lines after Bm (30 μ g/ml, overnight) treatment. Additionally, the colocalization experiment also confirmed that Bm-induced γ H2AX foci were higher and prolonged (up to 24 h) in DDX41 KO cells than WT cells (**Figure 16A, B, D, and E**). On average, 10-12 foci/cell were persistent in KO cell lines, whereas 1-2 foci/cell were detectable in WT cells after 24 h of DNA damage treatment (**Figure 16C and F**). Collectively, these results suggested that DDX41 might colocalize at DSB sites and participate in the DNA repair process along with other proteins and factors in the DNA-insulted cells.



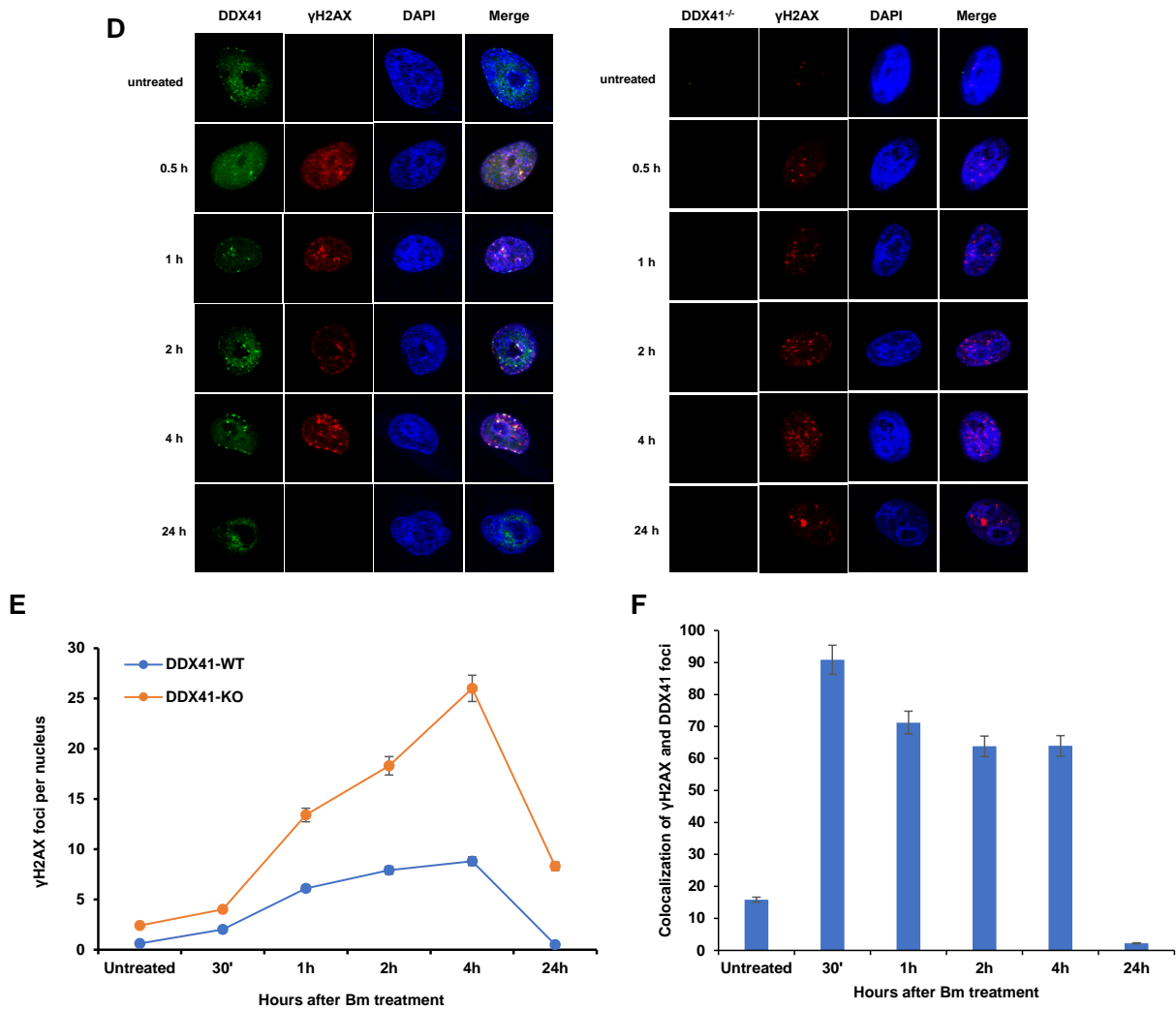


Figure 16. DDX41 colocalizes with γ H2AX upon DNA insult. (A and D) Immunofluorescence staining of WT and DDX41^{-/-} (A) HT1080 and (D) HeLa cells with γ H2AX (red), DDX41 (green), and DAPI (blue) at various time points (0, 0.5, 1, 2, 4, and 24 h) after Bm (30 μ g/mL, overnight) treatment in three biological replicates. Quantitative analysis of the percentage colocalization of γ H2AX and DDX41 after Bm treatment in HT1080 (B) and HeLa (E) cells. Quantitative analysis of the number of γ H2AX foci after Bm treatment in HT1080 (C) and HeLa (F) cells. The number of cells analyzed per channel are 100. Error bars indicate mean \pm SD of cells per channel. UT, untreated; Bm, bleomycin.

4.5 Comet assay validates higher and prolonged DNA damage in DDX41-KO cells

Next, I performed an alkaline comet assay to measure the DNA breaks formed in cells expressing or lacking DDX41 (Langie et al., 2015). Here, I estimated the amount of migrated DNA, revealed by the tail length indicating the extent of DNA damage in the cells (**Figure 17**). The untreated cells of both WT and DDX41-KO did not show any migration. However, DDX41-KO cells displayed higher migration than WT cells over 4 h after Bm treatment. In addition, this migration extended up

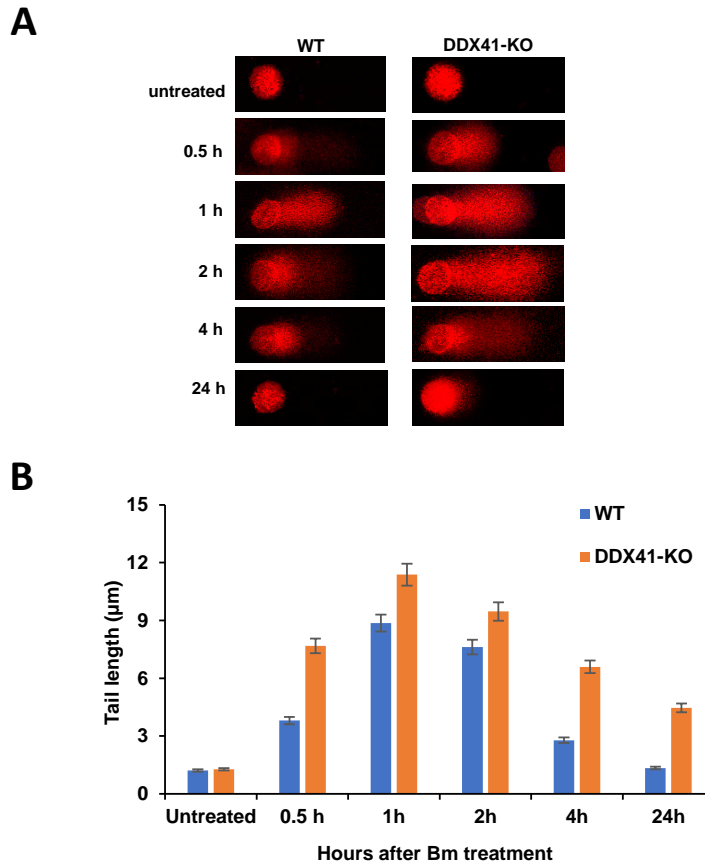


Figure 17. **Alkaline comet assay showing higher and prolonged DNA migration in DDX41-KO HT1080 cells.** (A) Staining DNA in WT and DDX41^{-/-} cells at various time points (0/untreated, 0.5, 1, 2, 4, and 24 h) after Bm (30 μg/ml, overnight) treatment **in three biological replicates** (B) Quantitative analysis of tail length (μm) (indicative of the damaged DNA) of the cells shown in panel A. Migration of comet tails of random 50 cells were measured by ImageJ program. Error bars indicate mean ± SD of cells per channel. UT, untreated; Bm, bleomycin.

to 24 h in KO cells, whereas this was resolved after 4 h of DNA damage in the WT cells (**Figure 17A and B**). The analysis of the tail length for WT cells demonstrates approximately 40-80% of the relative DNA migration, whereas, in DDX41-KO, this range grew to 80-100% over the time course after the Bm treatment. Moreover, after 24 hr of the DNA insult, the damaged DNA was recovered in the WT cells, i.e., the DNA migration was negligible, around 13%. In contrast, the DDX41-KO still showed significant unrecovered DNA damage, revealed by around 40% of the tail length (**Figure 17B**). Hence, the results of the alkaline comet assays were consistent with my previous Western blot and immunofluorescence results and confirmed that greater and persistent DNA damage occurs in DDX41 KO cells.

4.6 DDX41 promotes RNA:DNA hybrid clearance

Preliminary data in Dr. Wu's lab has shown that DDX41 protein binds and unwinds RNA:DNA hybrids *in vitro*, a part of the R-loop structure. Based on this, I wanted to determine the role of DDX41 in R-loop formation in cells. To address this, I performed dot blot assays for various time points (i.e., 0.25, 0.5, 1, 2, and 4 h) after DNA damage treatment in WT and DDX41-KO HT1080 cells. I used three different DNA damages, i.e., UV, HU, and Bm, and deciphered the levels of R-loops by dot blotting. To determine the R-loop structures in the cells and on the membrane, the antibody S9.6 was used. S9.6 is a gold standard antibody that is used to specifically detect RNA:DNA hybrids (Smolka et al., 2021). RNase H was used to degrade the RNA molecules in the RNA:DNA hybrid, serving as a negative control in the experiment. Furthermore, dsDNA was used as a loading control. As expected, RNase H degraded all R-loops in our samples. DDX41-KO cells treated with Bm (Figure 18A), UV (Figure 18B), or HU (Figure 18C) exhibited higher and prolonged R-loop structures, when compared to the wild type that resolves. To elaborate, upon Bm, UV, and HU treatment, both HT1080 WT and DDX41-KO cells exhibited R-loop formation. Here, in both the cell lines R-loops began to appear early in 0.25 h and increased by 0.5 h after the treatment. The WT cells resolved these hybrid structures within 1 h, on the other hand, in DDX41-KO cells these hybrids increased and accumulated up to 2 h and started delayed resolving at 4 h (Figure 18A). These results led me to conclude that DDX41 contributes to resolve R-loops.

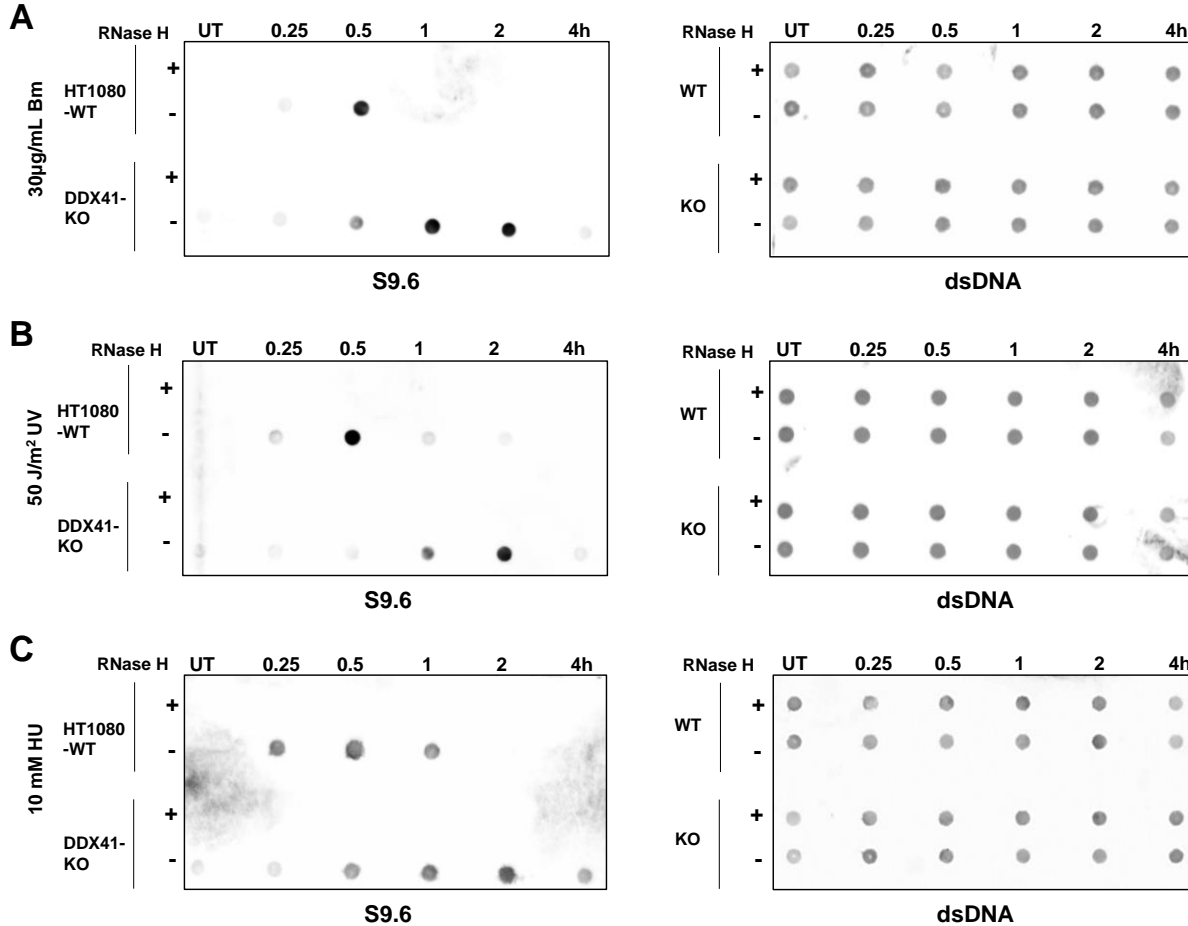


Figure 18. Dot blot assays showing increased and prolonged RNA:DNA hybrids in DDX41-KO cells. Dot blot assays using S9.6 antibody to detect R-loops in HT1080 (WT and DDX41-KO) cells at different time points (UT, 0.25, 0.5, 1, 2, and 4 h) after Bm (30 µg/mL, 16 h, **A**), UV (50 J/m², **B**), and HU (10 mM for 16 h, **C**) treatments in three biological replicates. dsDNA antibody was used as a loading control. UT, untreated.

4.7 Purification of DDX41-WT and DDX41-R525H mutant proteins

To investigate the molecular pathogenesis of patient mutation R525H, I expressed and purified DDX41-WT and DDX41-R525H proteins. After the Ni-NTA purification, DDX41-WT (**Figure 19A-C**) and DDX41-R525H (**Figure 19-F**) were purified to near homogeneity using size-exclusion chromatography on a Sephacryl S-300 HR 16/60 column. According to the molecular weight standards used to calibrate the size exclusion column, the molecular mass of the major peak, peak 2, was close to the monomeric form of DDX41 (69.8 kDa), while

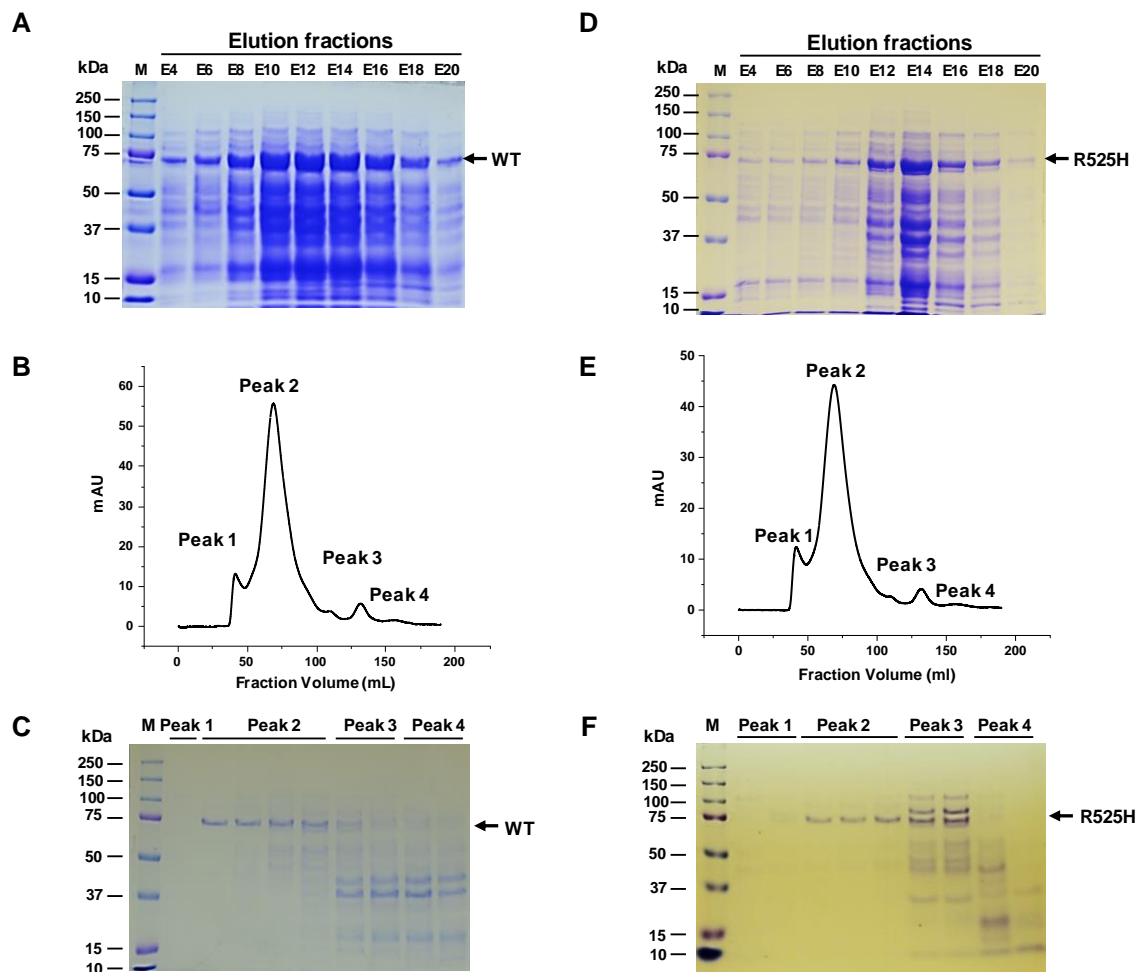


Figure 19. Purification of DDX41 proteins. (A and D) SDS-PAGE analysis of the eluted DDX41-WT (A) and DDX41-R525H (D) fractions from a Ni-NTA column. M, marker. Eluted fractions 4-20 are shown. (B and E) Chromatographic profiles of recombinant DDX41-WT (B) and DDX41-R525H (E) proteins eluted from a Sephacryl S-300 HR column. Four peaks are indicated. (C and F) SDS-PAGE analysis of the peaks shown in B and E.

peak 1 was in the void volume, likely aggregated form of the DDX41 proteins and peaks 3 and 4 were probably degraded fractions. Thus, I collected the monomer fractions of both DDX41 variants for the subsequent biochemical assays.

4.8 R525H protein has reduced unwinding activity but retains normal strand annealing activity

With the two purified proteins above, I wanted to check if DDX41 unwinds RNA:DNA hybrid, which mirrors the R-loop structure. I performed the helicase assay of both WT and R525H

mutant proteins. Compared with the DDX41-WT, a significant decrease in the unwinding activity was observed in the mutant (**Figure 20A**), indicating the mutant reduced its

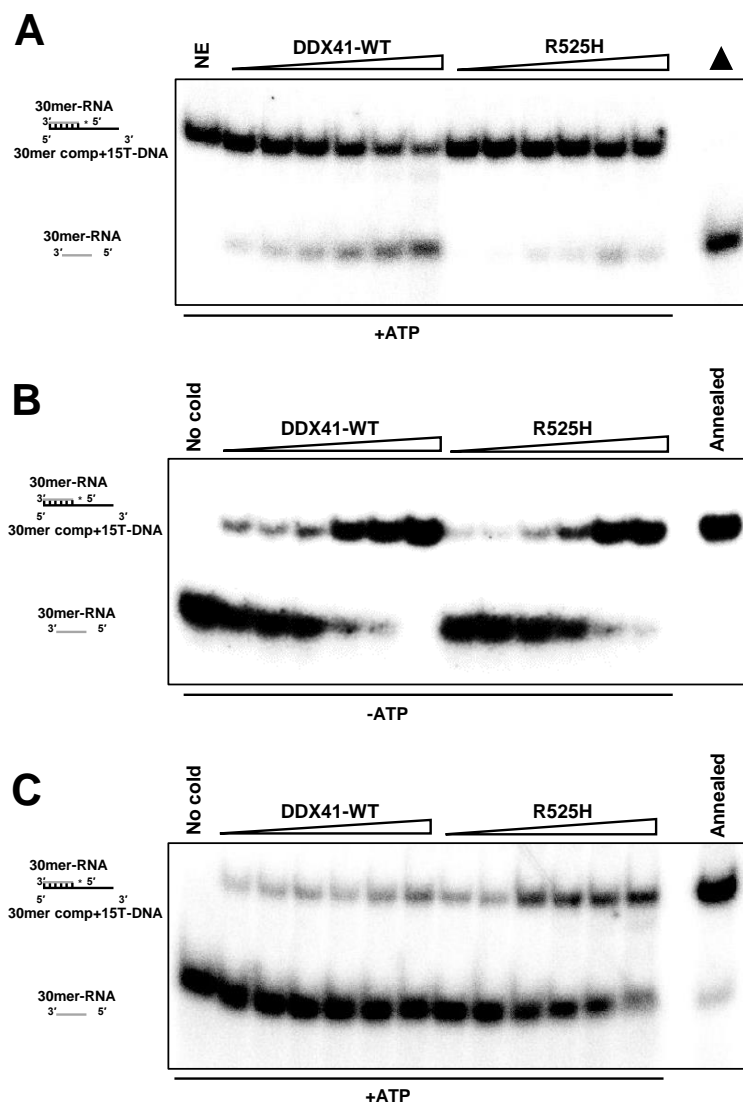


Figure 20. Helicase and annealing activities of DDX41-WT and R525H mutant proteins on RNA: DNA hybrids. (A) A representative image of helicase reactions performed by incubating 0.5 nM of 3' tail 30-mer RNA and 45-mer complementary DNA hybrid substrate with increasing protein concentration (0.09, 0.18, 0.37, 0.75, 1.5, 3.0 μ M) at 37 $^{\circ}$ C for 15 min. (B and C) Representative images of strand annealing assays performed by incubating 0.5 nM of P- labelled 30-mer RNA and 0.5 nM of unlabeled 45-mer complementary DNA with increasing protein concentration (0.09, 0.18, 0.37, 0.75, 1.5, 3.0 μ M) at 37 $^{\circ}$ C for 15 min, without ATP (B) and with ATP (C). DNA is represented with a black line and RNA with a grey line. NE, no enzyme; No cold, no unlabeled complementary strand; Filled triangle, heat-denatured substrate control.

helicase activity. Further, I compared strand annealing activities of both WT and mutant, and no significant difference was found (**Figure 20B**). Interestingly, when the same annealing assay reaction (shown in Figure B) was performed in the presence of ATP, WT unwound the RNA:DNA hybrid, whereas R525H failed to do so. Hence, there were accumulated RNA:DNA hybrids in the mutant reactions (**Figure 20C**). So, I found DDX41-WT efficiently performs both helicase and annealing activity whereas DDX41-R525H mutant has reduced helicase activity but retained annealing activity. To elaborate, in the presence of ATP, R525H showed significantly high strand annealing when compared to WT; whereas, although WT exhibited annealing activity but simultaneously it unwound the RNA:DNA hybrid formed after the annealing process. This led us to conclude that WT binds to the RNA:DNA hybrid structures in the R-loops and then resolves the structure for further DDR processing. In summary, these data suggested that DDX41-WT resolves RNA:DNA hybrids, whereas R525H tends to accumulate it.

4.9 DDX41 co-localizes with R-loops

Since DDX41 negatively regulates R-loops formation, next, I wanted to investigate whether DDX41 directly interacts with R-loop structures, i.e., colocalizes with RNA:DNA hybrids. Here, I used RNase H again to specifically degrade RNA:DNA hybrids. I treated WT (**Figure 21A**) and DDX41-KO cells (**Figure 21B**) with Bm to observe S9.6-foci formation and colocalization in WT cells (if any) at 1 h. Intriguingly, I found that DDX41 colocalized with R-loops/S9.6 in the absence of RNase H (**Figure 21A and D**), indicating its direct interaction with R-loops in the nucleus. Moreover, the number of S9.6-foci was higher in KO cells (average 10 foci/cell) than in WT cells (4 foci/cell) after 1 h of Bm treatment (**Figure 21B and C**). There is prominent S9.6 signal in nucleoli and cytoplasm that presumably largely originates from ribosomal RNA (Smolka et al., 2021). Strikingly, these results correspond with our *in vitro* (**Figure 20**) and dot blot data (**Figure 18**). Therefore, I conclude that DDX41 colocalizes with R-loops and resolves the formation.

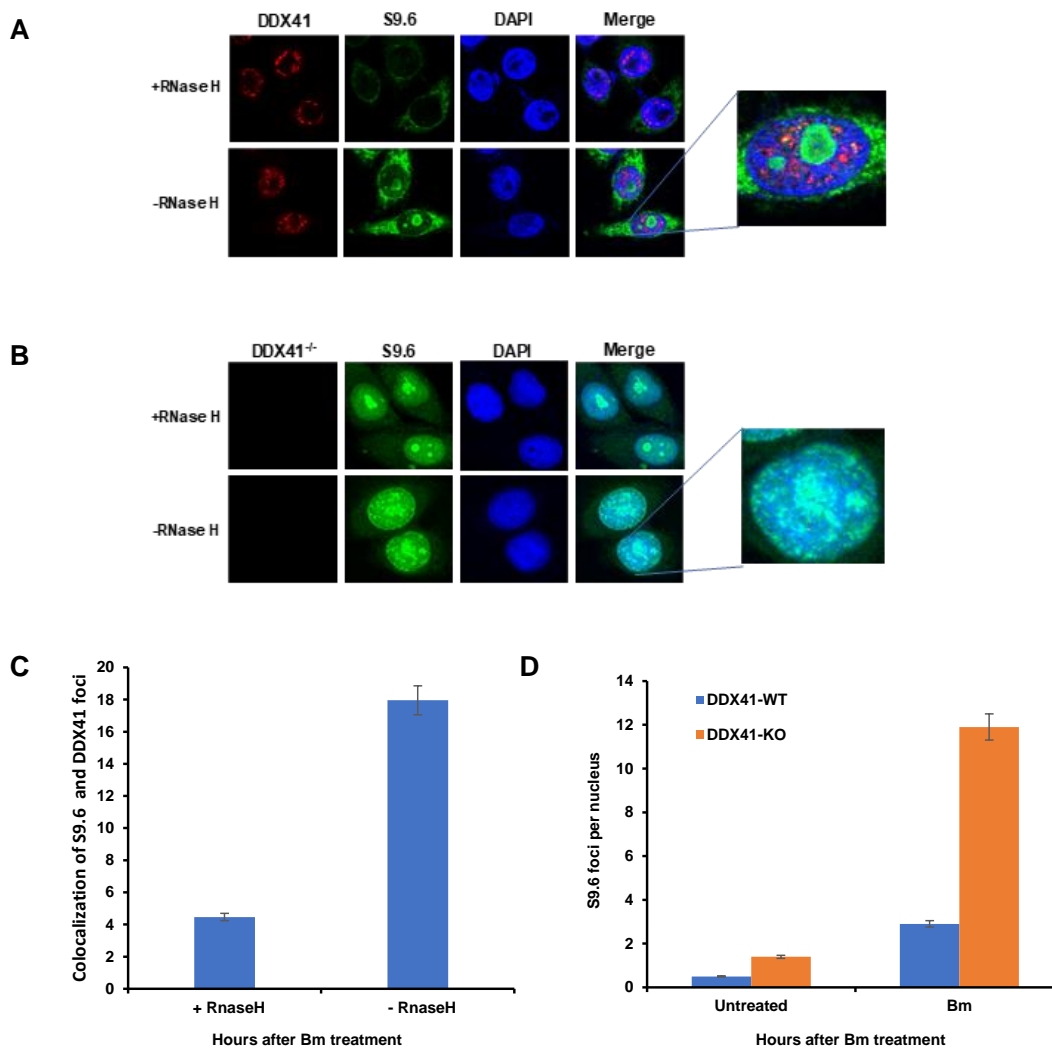


Figure 21. DDX41 colocalizes with R-loops upon DNA insult. Immunofluorescence staining of WT (**A**) and DDX41-KO (**B**) HT1080 cells at 1 h after Bm (30 μ g/mL, 16 h) treatment, with and without RNase H treatment in three biological replicates. (**C**) Quantitative analysis of the percentage colocalization of S9.6 and DDX41 shown in panel A. (**D**) Quantitative analysis of S9.6 foci in WT and DDX41-KO HT1080 cells without or with Bm treatment (16 h post 30 μ g/mL). The number of cells analyzed are 100 per channel. Error bars indicate mean \pm SD of cells per channel. UT, untreated; Bm, bleomycin.

4.10 DDX41 colocalizes with transcriptional machinery

Because R-loops are frequently formed during transcription (Belotserkovskii et al., 2018), I used RNA polymerase II antibody and performed confocal imaging to investigate whether

DDX41 colocalizes with the transcriptional machinery and if there is any difference in colocalization before and after DNA damage. There was no colocalization between DDX41 and RNA Pol II in the absence of DNA damage. After the bleomycin treatment, around 70 % of DDX41 colocalized with RNA pol II at 1 h, furthermore, there was 1.5 times increase in RNA Pol II in HT1080 cells (**Figure 22**). Possibly, upon the bleomycin treatment, RNA Pol II accumulated at the DSB sites due to halted passage at the lesion site, where DDX41 is recruited to resolve R-loops.

Besides DDX41's potential role in DSB-induced R-loops metabolism, I wanted to know whether transcription-replication collisions cause the increased DSBs in DDX41

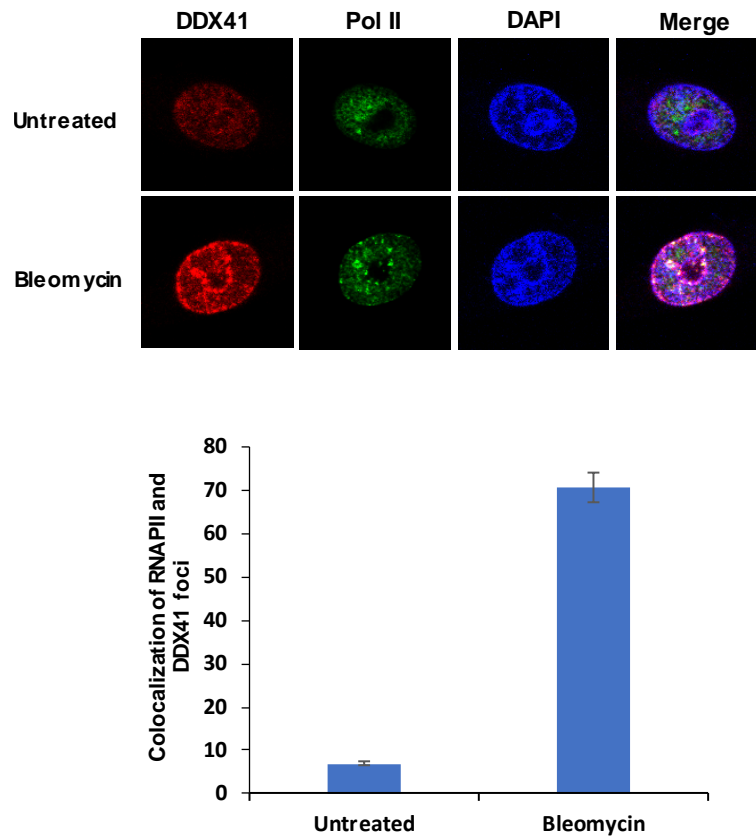


Figure 22. DDX41 colocalizes with RNA Pol II upon bleomycin treatment. Immunofluorescence staining of HT1080 cells after bleomycin (30 $\mu\text{g/mL}$, overnight) treatment in three biological replicates and quantitative analysis of percentage colocalization of RNA Pol II and DDX41 shown in panel A. The number of cells analyzed are 100 per channel. Error bars indicate mean \pm SD of cells per channel. UT, untreated.

depletion cells. I have sent our WT and DDX41-KO cell lines to Dr. Peter Sterling's lab at UBC, who will measure the transcription–replication collision events as PLA (Proximity Ligation Assay) signals between PCNA (Proliferating cell nuclear antigen, for replication) and RNA Pol II (transcription).

4.11 Patient mutant R525H exhibits diminished DDR and increased DNA damage when reconstituted in wildtype or DDX41 KO cells

To explore the potential cellular pathogenesis of the R525H mutant, I overexpressed the DDX41-WT or DDX41-R525H genes in both WT and DDX41-KO HT1080 cell lines. After Bm, UV, and HU treatment, I examined the relative abundance of the DDR-related proteins and R-loops status by Western and dot blotting respectively. Like the Western blot results of DDX41-KO HT1080 cells (**Figure 12**), both WT and DDX41 KO backgrounds overexpressing the R525H mutant showed significantly higher γ H2AX and 53BP1 levels along with delayed and diminished DNA repair levels, indicated by pBRCA1 protein level after Bm (**Figure 23A**), UV (**Figure 23B**) and HU (**Figure 23C**) treatment. DSB, indicated by γ H2AX and 53BP1 levels, was resolved within 2-4 h after the DNA damage in WT overexpression cell lines, but it took 8-24 h in DDX41-R525H overexpressing cells, both in WT and DDX41 KO cell background. Moreover, the phosphorylation of BRCA1 protein was earlier, around 1-2 h, in WT overexpressed cells than in DDX41-R525H overexpressed cells, i.e., around 4-8 h after the DNA damage.

In addition, dot blot assays also showed that R525H had elevated and persistent RNA:DNA hybrid levels (until 4 h) when compared to WT (until 0.5-1 h) after Bm (**Figure 24A and 25A**), UV(**Figure 24B and 25B**), or HU (**Figure 24C and 25C**) treatment in DDX41-KO (**Figure 24**) and WT (**Figure 25**) cell background. Thus, I concluded that the patient mutant R525H dysregulates DDR and DNA repair pathways, which may explain the underlying pathogenesis in MDS and AML patients that harbor the R525H mutation.

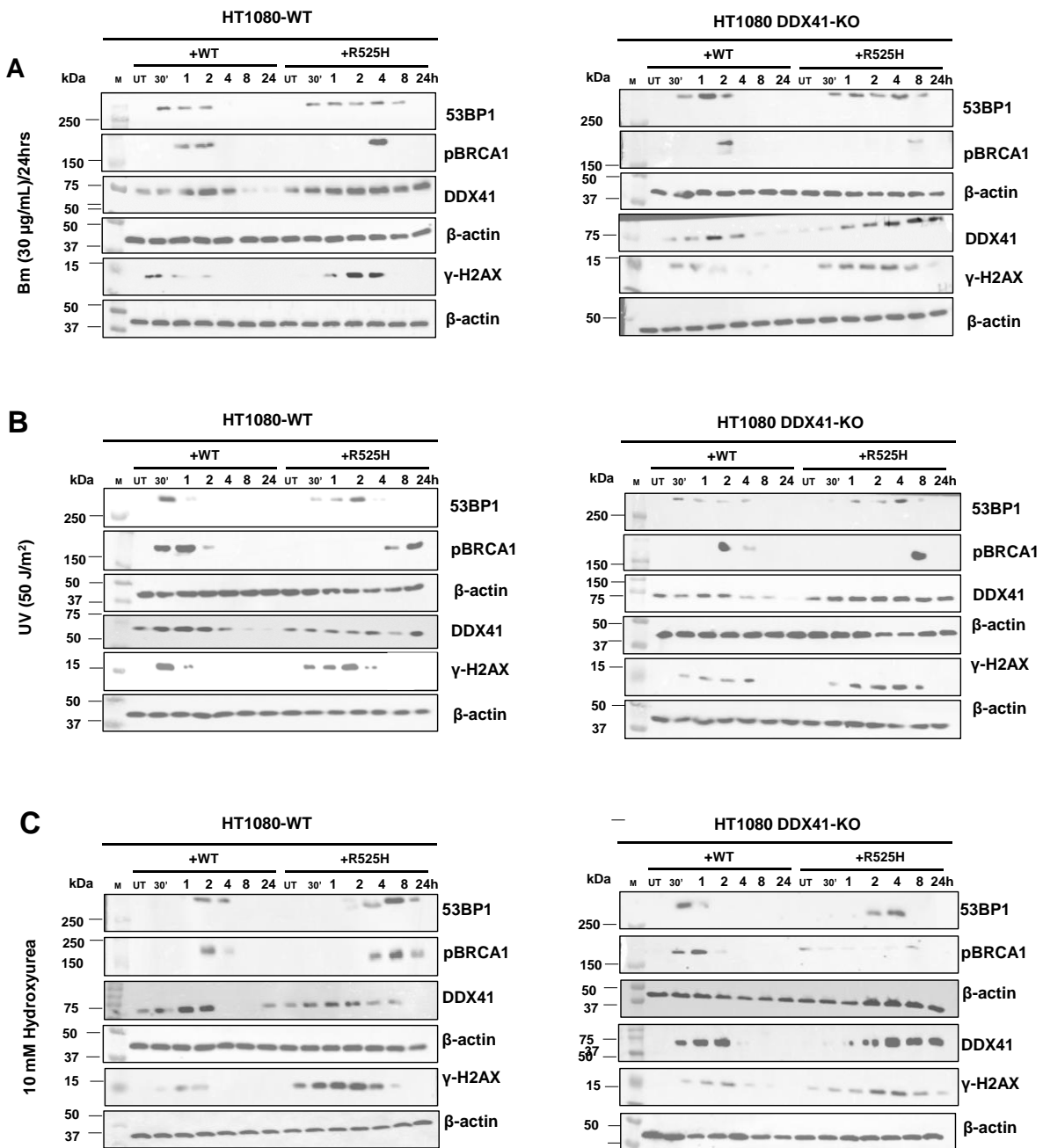


Figure 23. Overexpression of the R525H mutant exhibits prolonged DNA damage and diminishes repair. Western blot assays of some DDR proteins in DDX41-WT or DDX41-R525H overexpressed WT and DDX41-KO HT1080 cell lines after Bm (30 µg/mL for 16 h, **A**), UV (50 J/m², **B**), and HU (10 mM, 16 h, **C**) at various time points (0.5, 1, 2, 4, 8, and 24 h) in three biological replicates. β-actin serves as a loading control. UT, untreated.

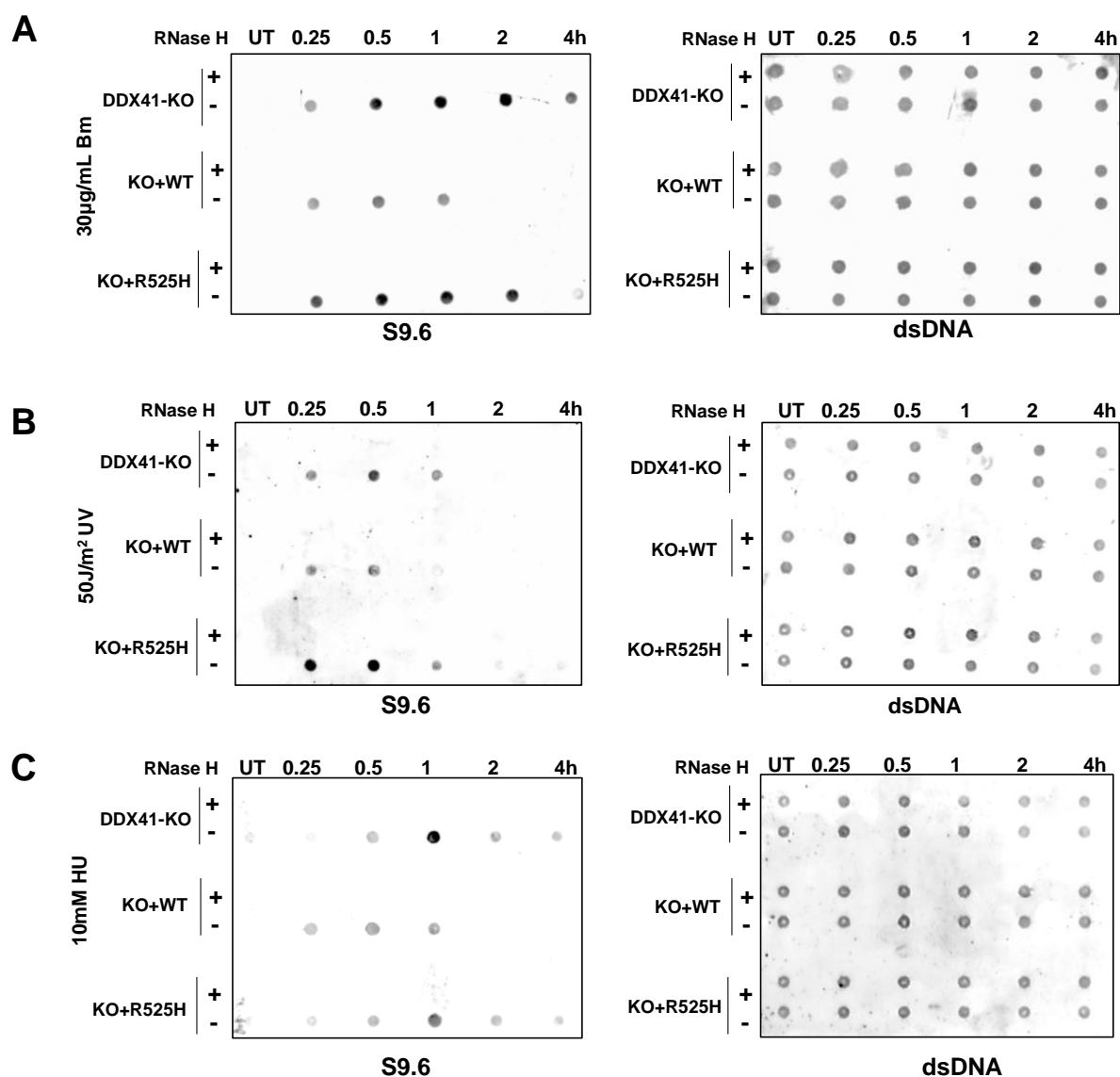


Figure 24. Overexpression of the R525H mutant perturbs RNA:DNA hybrids in DDX41-KO HT1080 cells. Dot blot assay of HT1080 DDX41-KO cells and DDX41-KO cells expressing DDX41-WT or DDX41-R525H genes at different time points (UT, 0.25, 0.5, 1, 2, and 4 h) after Bm (30 µg/mL, 16 h, **A**), UV (50J/m², **B**), and HU (10 mM, 16 h, **C**) treatments in three biological replicates. S9.6 antibody was used to detect R-loops and dsDNA antibody as a loading control. UT, untreated.

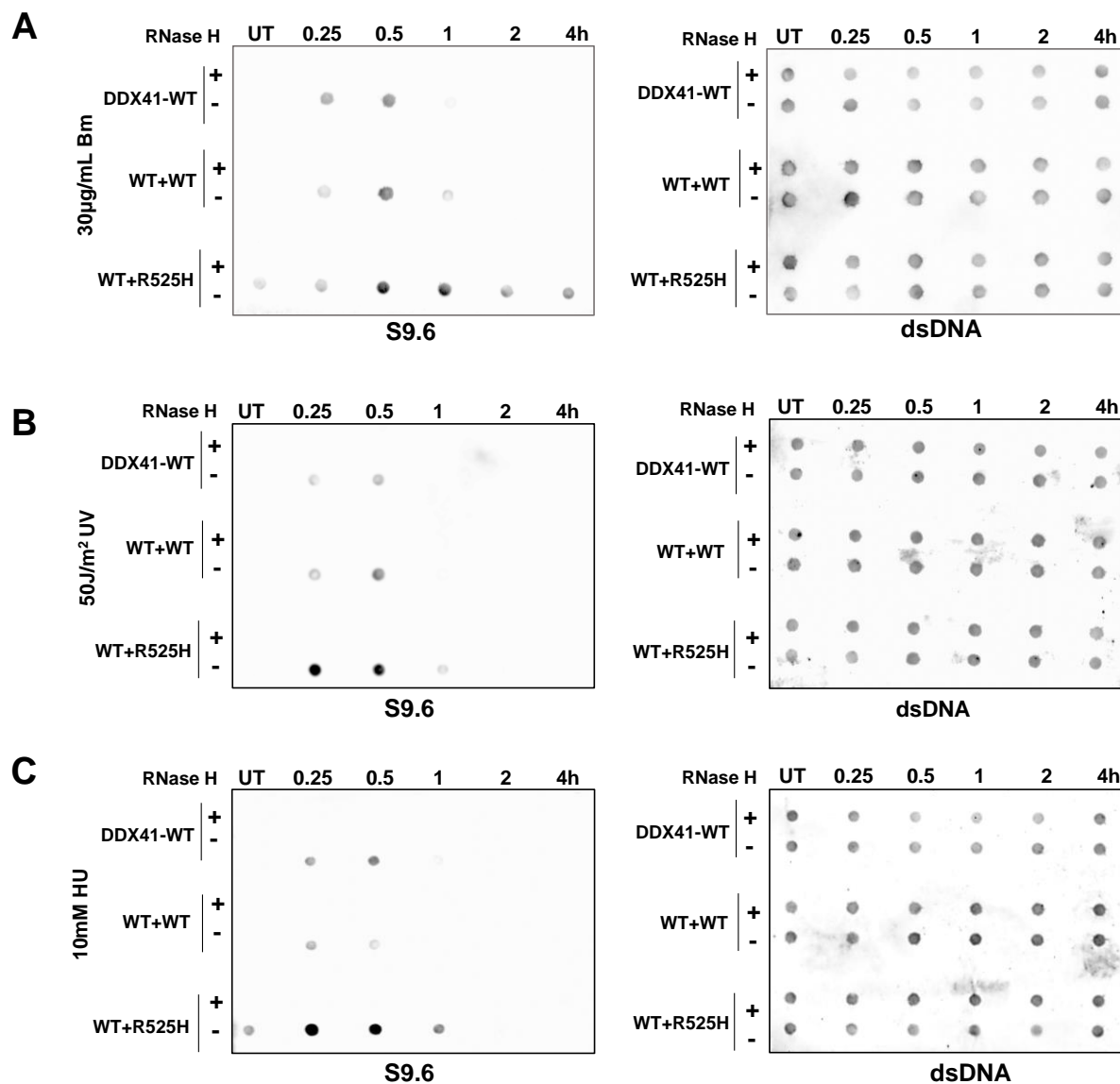


Figure 25. Overexpression of the R525H mutant perturbs RNA:DNA hybrids in WT HT1080 cells. Dot blot assay of HT1080 WT cells and WT cells expressing DDX41-WT or DDX41-R525H genes at different time points (UT, 0.25, 0.5, 1, 2, and 4 h) after Bm (30 µg/mL, 16 h, A), UV (50J/m², B), and HU (10 mM for 16 h, C) treatments in three biological replicates. S9.6 antibody was used to detect R-loops and dsDNA antibody as a loading control. UT, untreated.

5 DISCUSSION

5.1 DDX41 is an emerging DDR-associated protein

DDX41 is well known for its role as a host intracellular DNA sensor against DNA virus infection (Zhang et al., 2011c). Both germline and acquired somatic mutations of DDX41 stimulate the development of MDS and AML (Abou Dalle et al., 2020; Cardoso et al., 2016; Ding et al., 2012; Hosono, 2019; Saygin and Godley, 2021). To investigate the possible mechanism involved in the pathogenesis of MDS/AML, extensive research has been taking place, and now the attention has been driven to DNA damage and alteration of the DDR, which are critical features of deregulation of genomic stability (Boehrer et al., 2009; Cavelier et al., 2009; Kefala et al., 2013; Popp et al., 2017; Satoh et al., 2012). DEAD-box proteins, such as DDX1 (Li et al., 2008; Zhang et al., 2011a) and DDX3 (Sun et al., 2013; Szappanos et al., 2018), which have well-known functions in the innate immune response, are also intriguingly involved in DDR. On the other hand, several DDR genes, such as ATM (Grosjean-Raillard et al., 2009), XPD and XRCC1 (Joshi et al., 2016), BRCC (Meyer et al., 2020), and HLTF (Takaoka et al., 2019) have been associated with MDS/AML.

DEAD-box helicases are usually involved in RNA metabolism, such as mRNA splicing, translation initiation, ribosome biogenesis, and RNA degradation (Valentini and Linder, 2021); however, so far more than a dozen of DEAD-box helicases has been identified in DNA repair. They are DDX1 (Kim et al., 2020; Li et al., 2016a), DDX3 (Chan et al., 2019), DDX5 (Kim et al., 2020), DDX6 (Bergkessel and Reese, 2004), DDX11 (Abe et al., 2018; Guo et al., 2015; Shah et al., 2013), DDX17 (Adamson et al., 2012; Koo et al., 2015; Li et al., 2016b), DDX19 (Hodroj et al., 2017), DDX21 (Kim et al., 2020), DDX23 (Sridhara et al., 2017), DDX47 (Okamoto et al., 2019), DDX54 (Milek et al., 2017), DDX56 (Kouyama et al., 2019), DDX58 (Malachin et al., 2017; Ranoa et al., 2016), and DHX9 (Chakraborty et al., 2018b; Cristini et al., 2018; Kim et al., 2020). To briefly mention, DDX5, also known as p68, localizes in DNA damage sites (Kahlina et al., 2004). DHX9, also known as RNA helicase A, is associated directly with DNA repair marker γ -H2AX after actinomycin D treatment in HeLa cells (Mischo et al., 2011; Mischo et al., 2005). DDX1 is found to facilitate the removal of RNA and homologous recombination at DSBs (Li et al., 2016a). Biomedical and genetic approaches have identified several other DEAD-box helicases, though their exact roles in DDR remain under

further investigation.

The potential role of DDX41 in DDR is just emerging. Recently, in a zebrafish model, a study revealed the importance of DDX41 in erythropoiesis regulation as deficiency and mutations in DDX41 exhibits ATM- and ATR-mediated cell arrest triggered by DNA damage, genomic stress, and misexpression and alternative splicing of genes related to cell cycle (Weinreb et al., 2020). My data provide evidence that DDX41 is another DEAD-box protein that functions in DDR: it assists DNA repair by resolving R-loop structures and precludes genome instability.

5.2 DDX41 is required to facilitate DNA repair

My work shows that with the loss of DDX41, the DNA damage measured by γ -H2AX and 53BP1 levels becomes extensive and longer in duration; in addition, the DNA repair signalling (phosphorylation of BRCA1 to pBRCA1) was diminished and occurred much later compared to wildtype cells. The DNA damage-induced γ -H2AX foci formation is a critical and primary cellular response to DSB, and the number of γ -H2AX foci decreases as DSBs are repaired (Redon et al., 2009). In our confocal imaging studies, γ -H2AX foci were resolved over the time course of 4 h after various types of DNA damage in wild type cells, whereas in DDX41 knockout cells these foci persisted until 24 h. DDX1 is recruited to the DNA damage sites, early at 30 min upon IR exposure and much later after 24 h of UV and cisplatin treatments, thus facilitating double strand break repair (Li et al., 2008). An alkaline comet assay was performed to assess the lesions that particularly signifies DSB. Intriguingly, the alkaline comet assays bolstered our Western blots and immunofluorescence results, i.e., the tail moment indicated by the damaged DNA migrated in the alkaline comet assay was longer in DDX41-KO (24 h) than WT cells (4 h).

The DDR factor 53BP1 marks the DNA damage site after γ -H2AX recruitment and regulates the choice between HR and NHEJ DNA repair pathways (Bunting et al., 2010; Chapman et al., 2013). Similarly, pBRCA1 chooses between HR and NHEJ repair pathways by deciding on various unified factors (Biehs et al., 2017). Interestingly, when 53BP1 and pBRCA1 localize collectively on the damage site, they promote for HR-mediated repair. Our Western blot results show the levels of 53BP1 and pBRCA1 largely mirrored each other, suggesting that DDX41 might be involved in HR-mediated DNA repair. Hence, the persistent and

higher DNA damage and delayed DNA repair protein levels in DDX41-knockout cells reflect the importance of DDX41 in DSB repair.

5.3 DDX41 resolves R-loops

The novel function of DDX41 in relation to R-loop structures is just emerging. Using zebrafish, it was found excess R-loop structures formed in DDX41 mutant animals, which deterred the equilibrium of hemogenic endothelium, hematopoietic stem and progenitor cells required for cellular fitness (Weinreb et al., 2021). Another research revealed DDX41 depletion leads to replication stress, DNA damage, inflammatory signaling and accumulation of R-loop structures, whereas this accumulation is opposed in the presence of DDX41 through unwinding RNA:DNA hybrids at gene promoter (Mosler et al., 2021). Correspondingly, our results showed that DDX41 protein unwinds RNA:DNA hybrids *in vitro*, suggesting its function in removing and regulating R-loops. I also performed *in vitro* annealing assays using DDX41-WT and R525H mutant proteins and deciphered that both could anneal the RNA: DNA hybrids without any significant difference. Excitingly, a significant difference was observed when this annealing assay was performed in the presence of ATP, whereas the DDX41 mutant showed accumulation of R-loop structures and WT was able to unwind the RNA:DNA hybrids. Herein, I concluded that DDX41 is required for resolving and regulating R-loops.

Compelling evidence has implicated several other DEAD-box proteins in the suppressing co-transcriptional R-loop structures, including DDX19, DDX21, DDX23, and DDX47. In 2017, DDX19 was discovered to utilize its helicase activity to resolve R-loops *in vitro* (Hodroj et al., 2017). In a similar study, the phosphorylation of DDX23 by a serine/arginine protein kinase 2 (SRPK2) was found to initiate a signal transduction cascade to help DDX23 resolve R-loops (Sridhara et al., 2017). Also, DDX21 was discovered to cooperate with SIRT7 to resolve R-loops and safeguard genome stability (Song et al., 2017). In 2018, DDX47 was added to the list as it interacts with Fanconi anemia protein, FANCD2, during mild replication stress to decrease the number of R-loops by reducing transcription–replication collisions (Okamoto et al., 2019). In 2019, DDX21, together with H3K27me3 demethylase JMJD3, was shown to resolve aberrant R-loops (Argaud et al., 2019). In 2020, DDX5 was reported to promote homologous recombination repair by acting as an R-loop-resolvase at the sites of DSBs in U2OS- 265 and DRGFP cells (Yu et al., 2020). Thus, a fundamental question is why do cells need so many R-loop helicases, even

in the same DEAD-box family? One possibility is that various DEAD-box helicases act distinctly on R-loops based on (1) type of R-loop, (2) different genome sites, and 3) cell or tissue-specific. Nevertheless, further studies are required.

5.4 DDX41: a potential co-transcriptional suppressor of transcription replication conflicts

Transcription-replication collisions (TRCs), formed due to unresolved R-loop structures, are another threat to genomic stability. Many reports suggest that some well-identified DNA repair proteins like BRCA1 and FANCD2 function as R-loop resolvases at the TRC locations, but the mechanism is still not explored well (Bhatia et al., 2014; Domínguez-Sánchez et al., 2011; Schwab et al., 2015). In my project, I anticipate that the absence of DDX41 (or in the presence of R525H mutant) might result in higher TRCs due to R-loop accumulation in the cells. Studies in this field have revealed the R-loop resolving helicase, Senataxin, moves along with the replication fork to promote the progression through RNA polymerase II transcribed genes (Alzu et al., 2012; Mischo et al., 2011; Skourti-Stathaki et al., 2011). Also, the DEAD-box helicase DDX39B is implicated in inhibiting R-loop accumulation during transcription and suppressing transcription-replication conflicts upon DNA damage (Pérez-Calero et al., 2020). To date, I have found the colocalization of DDX41 and RNA polymerase II, suggesting DDX41 might interact with the transcription machinery to help it release from the DSB-prone sites and maintain the genome integrity from pathological R-loop accumulation. Further, ongoing proximal ligation assays (PLA) will give me direct evidence, measuring the collision events between replication and transcriptional machinery, *i.e.*, PCNA and RNA Pol II, respectively, in WT and DDX41 KO cells after DNA insult.

5.5 Prospective molecular pathogenesis of MDS and AML R525H mutation

My study suggests that in the absence of DDX41, cells display elevated DSBs and delayed and diminished pBRCA1 activation. My work also identifies that DDX41 is a pivotal R-loop resolvase that facilitates the repair of DSBs. By performing dot blots after different DNA damage, I observed increased and accumulated R-loop structures in DDX41-R525H mutated cells, and significantly lower R-loops in the cells with the presence of DDX41-WT. Parallel to

my findings, another DEAD-box helicase, DDX5, is overexpressed in acute myeloid leukemia, counteracting DSB-related DNA deletions and repair defects by resolving persistent R-loops (Yu et al., 2020). My working model (**Figure 26**) supplements the exciting possibility of

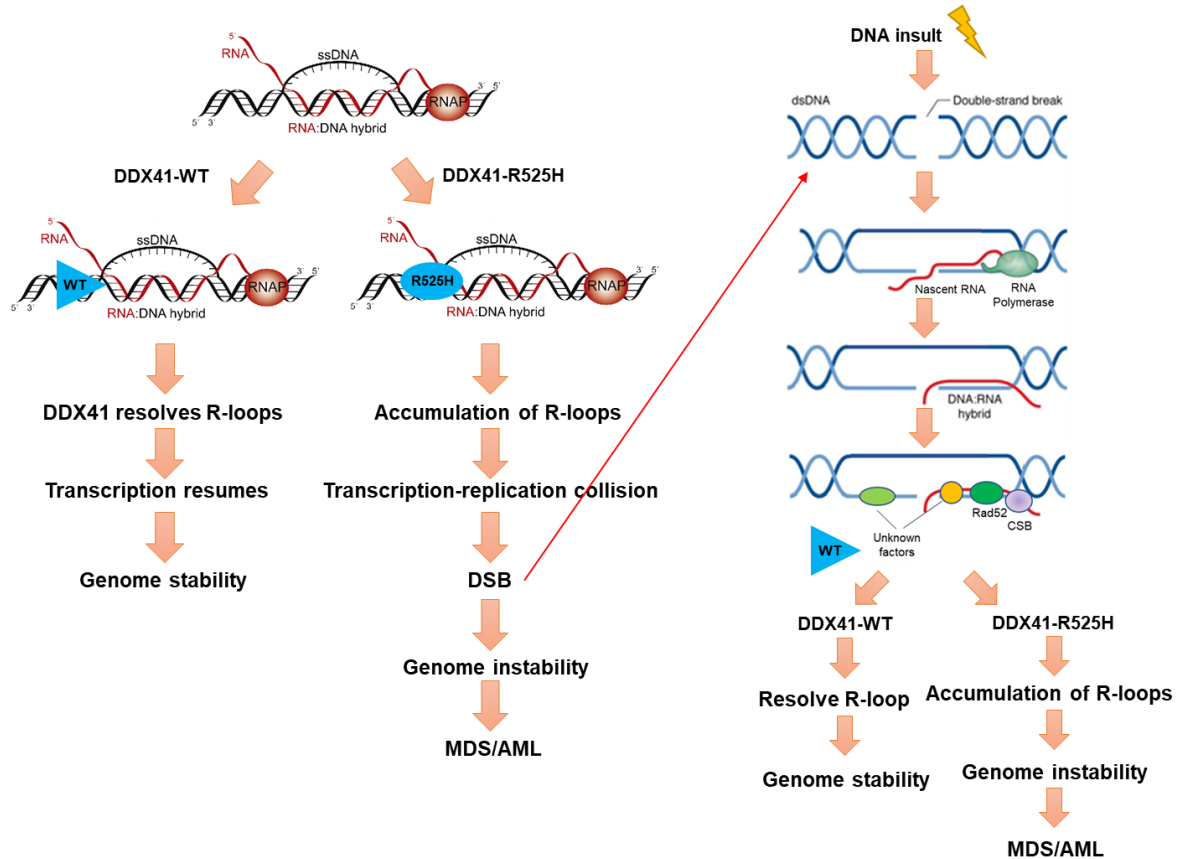


Figure 26. A proposed model of the role of DDX41 in co-transcriptional resolution of R-loops to maintain genome stability. RNAP, RNA polymerase; DSB, Double-strand breaks; WT, Wildtype; CSB, Cockayne syndrome protein B. Modified from (Hamperl and Cimprich, 2014; Hegazy et al., 2020).

DEAD-box helicases associated with R-loop and other aberrant structures. The model illustrates that DDX41-WT co-transcriptionally resolves any unwanted R-loop structure during transcription (**Figure 26, left**), whereas DDX41-R525H mutant fails to do that, hence, accumulates R-loops. The accumulation of R-loops results in transcription-replication collisions, DSBs, genome instability, and possibly MDS and AML. Similarly, another DEAD-box helicase, DDX39B, was recently determined to co-transcriptionally remove the pathological R-loop structures as its overexpression leads to hybrid suppression and restraining of the R-loop accumulation (Pérez-Calero et al., 2020). Alternatively (**Figure 26, right**), DSBs are formed

as a consequence of DNA damage (from any source) in the cells, and there is a tendency of R-loop accumulation at the site (Cristini et al., 2019). DDX41-WT and other DNA repair proteins like RAD52 and CSB resolve R-loops, whereas DDX41-R525H mutant fails, resulting in delayed DNA repair and possible MDS/AML. Correspondingly, DDX21 collaborates with SIRT7 to check any R-loop formation and maintain genome integrity (Song et al., 2017). Describing the collaboration above, both SIRT7 and DDX21 are related to Pol I and Pol II. In the wild type cells, DDX21-related hypoacetylation/deacetylation which is needed for unwinding of co-transcriptional R loops is maintained, where SIRT7-KO cells results in impaired helicase activity due to hyperacetylation of DDX21, resulting in stalled activity of Pol I and Pol II, R loop accumulation, and eventually DNA damage. Many other DEAD-box helicases such as DHX9, DDX5, DDX1, and DDX21 interact with ATAD5, a PCNA unloader, to diminish RNA:DNA hybrids under both standard and replication stress conditions (Kim et al., 2020).

It is exciting to apply this model of DDX41 to the pathogenesis of MDS and AML; however, additional pathways and clinical studies on how these pathways interplay and influence each other should be investigated to make the pathogenesis effect evident. Our work opens a window that illustrates how DDX41 associated regulation of DNA damage response pathway and R-loop formation are interlinked and possibly applicable to MDS and AML pathologies and other diseases.

6 CONCLUSIONS AND FUTURE WORK

6.1 Conclusions

DDX41 mutations are associated with hematologic malignancies, including MDS and AML; however, the molecular pathogenesis is unknown. Since dysfunction of DNA damage and repair is one of the molecular pathogenesis of MDS and AML, I explored the potential role of DDX41 in DNA repair. Using HT1080 and HeLa DDX41 KO cell lines, I found DDX41-depletion leads to increased DNA damage, indicated by increased and prolonged DNA double strand breaks and increased comet tail formation. Further, I found there were increased R-loops in DDX41 KO cells after DNA insult. DDX41 protein resolves R-loops *in vitro*, but the MDS and AML patient mutant R525H marginally unwinds R-loop structures. DDX41 co-

localized with DSB marker γ -H2AX and R-loop marker S9.6. When overexpressed in DDX41 KO or wildtype cell lines, R525H-expressing cells had significantly more R-loops than WT-expressing cells. Taken together, my results suggest DDX41 utilizes its unwinding activities to resolve R-loops at the DSB sites, which is critically important for DSB repair and genome stability, and dysregulation of this pathway might lead to MDS/AML.

Even though if the confocal microscope is perfectly tuned it is not adequate to ascertain whether two proteins tagged with fluorescent molecules are colocalizing. Therefore, more specific techniques like pull down assay and proximity ligation assay must be used for future studies that readily discriminate over specific antibodies or probes targeting different proteins.

6.2 Future work

To confirm the correlation of DNA damage and DDR with DDX41, I will investigate the colocalization of DDX41 with 53BP1, another DSB marker, since both γ -H2AX and 53BP1 can be visualized as foci at the sites of DSB by confocal microscopy (Rothkamm et al., 2015). Further, our Western blots and dot blots showed there were elevated γ -H2AX and R-loop levels in R525H-expressed cell lines; additional immunofluorescence experiments will yield perceptive information about whether R525H mutation has a direct localization defect. Hence, the colocalization of GFP-tagged DDX41-WT and GFP-tagged DDX41-R525H with DSB marker γ H2AX/53BP1 could be examined and compared to produce further details on whether and how mutations in DDX41 alter the DDR pathway, particularly the initial steps of DNA damage responses.

Although various experiments led us to conclude the importance of DDX41 in DSB repair by resolving R-loops, whether DDX41 assists in the HR repair pathway remains unknown. To address this, I-SceI endonuclease that generates a DSB in the cassette of the dysfunctional GFP gene fragment in DR-GFP U2OS cell line (Gunn and Stark, 2012) can be used to compare HR efficiency in WT and siRNA DDX41 knockdown cell lines. This DSB formed by I-SceI can be repaired via HR and the downstream GFP can be used as the readout to measure HR efficacy. Lastly, the investigation of transcription and replication collisions (TRCs) status in WT vs. DDX41-KO cell lines, DDX41-WT vs. DDX41-R525H-expressed cell lines using proximity ligation assay (collaborating with Dr. Peter Stirling at UBC) could also be informative to interpret the potential role of DDX41 in the molecular pathogenesis of MDS/AML.

7. REFERENCES

- Abe, T., Ooka, M., Kawasumi, R., Miyata, K., Takata, M., Hirota, K., and Branzei, D. (2018). Warsaw breakage syndrome DDX11 helicase acts jointly with RAD17 in the repair of bulky lesions and replication through abasic sites. *Proc Natl Acad Sci U S A*. *115*, 8412-8417.
- Abou Dalle, I., Kantarjian, H., Bannon, S.A., Kanagal-Shamanna, R., Routbort, M., Patel, K.P., Hu, S., Bhalla, K., Garcia-Manero, G., and DiNardo, C.D. (2020). Successful lenalidomide treatment in high risk myelodysplastic syndrome with germline DDX41 mutation. *Am. J. Hematol.* *95*, 227-229.
- Adamson, B., Smogorzewska, A., Sigoillot, F.D., King, R.W., and Elledge, S.J. (2012). A genome-wide homologous recombination screen identifies the RNA-binding protein RBMX as a component of the DNA-damage response. *Nat. Cell Biol.* *14*, 318-328.
- Agafonov, D.E., Deckert, J., Wolf, E., Odenwälder, P., Bessonov, S., Will, C.L., Urlaub, H., and Lührmann, R. (2011). Semiquantitative proteomic analysis of the human spliceosome via a novel two-dimensional gel electrophoresis method. *Mol. Cell. Biol.* *31*, 2667-2682.
- Aguilera, A., and García-Muse, T. (2012). R loops: from transcription byproducts to threats to genome stability. *Mol. Cell* *46*, 115-124.
- Almeida, A.M., and Ramos, F. (2016). Acute myeloid leukemia in the older adults. *Leuk. Res. Rep.* *6*, 1-7.
- Alzu, A., Bermejo, R., Begnis, M., Lucca, C., Piccini, D., Carotenuto, W., Saponaro, M., Brambati, A., Cocito, A., and Foiani, M. (2012). Senataxin associates with replication forks to protect fork integrity across RNA-polymerase-II-transcribed genes. *Cell* *151*, 835-846.
- Anderson, J.S.J., and Parker, R. (1998). The 3' to 5' degradation of yeast mRNAs is a general mechanism for mRNA turnover that requires the SKI2 DEVH box protein and 3' to 5' exonucleases of the exosome complex. *EMBO J.* *17*, 1497-1506.
- Argaud, D., Boulanger, M.-C., Chignon, A., Mkannez, G., and Mathieu, P. (2019). Enhancer-mediated enrichment of interacting JMJD3–DDX21 to ENPP2 locus prevents R-loop formation and promotes transcription. *Nucleic Acids Res.* *47*, 8424-8438.
- Auboeuf, D., Höning, A., Berget, S.M., and O'Malley, B.W. (2002). Coordinate regulation of transcription and splicing by steroid receptor coregulators. *Science* *298*, 416-419.
- Ball, H.L., Myers, J.S., and Cortez, D. (2005). ATRIP binding to replication protein A-single-stranded DNA promotes ATR–ATRIP localization but is dispensable for Chk1 phosphorylation. *Mol. Biol. Cell.* *16*, 2372-2381.
- Basu, U., Meng, F.-L., Keim, C., Grinstein, V., Pefanis, E., Eccleston, J., Zhang, T., Myers, D., Wasserman, C.R., and Wesemann, D.R. (2011). The RNA exosome targets the AID cytidine

deaminase to both strands of transcribed duplex DNA substrates. *Cell* 144, 353-363.

Bejar, R., Stevenson, K., Abdel-Wahab, O., Galili, N., Nilsson, B., Garcia-Manero, G., Kantarjian, H., Raza, A., Levine, R.L., and Neuberg, D. (2011). Clinical effect of point mutations in myelodysplastic syndromes. *NEJM*. 364, 2496-2506.

Belotserkovskii, B.P., Tornaletti, S., D'Souza, A.D., and Hanawalt, P.C. (2018). R-loop generation during transcription: Formation, processing and cellular outcomes. *DNA Repair* 71, 69-81.

Bergkessel, M., and Reese, J.C. (2004). An Essential Role for the *Saccharomyces cerevisiae* DEAD-Box Helicase DHH1 in G1/S DNA-Damage Checkpoint Recovery. *Genetics* 167, 21-33.

Bernstein, K.A., Gangloff, S., and Rothstein, R. (2010). The RecQ DNA helicases in DNA repair. *Annu. Rev. Genet.* 44, 393-417.

Bessonov, S., Anokhina, M., Will, C.L., Urlaub, H., and Lührmann, R. (2008). Isolation of an active step I spliceosome and composition of its RNP core. *Nature* 452, 846-850.

Bhatia, V., Barroso, S.I., García-Rubio, M.L., Tumini, E., Herrera-Moyano, E., and Aguilera, A. (2014). BRCA2 prevents R-loop accumulation and associates with TREX-2 mRNA export factor PCID2. *Nature* 511, 362-365.

Biehs, R., Steinlage, M., Barton, O., Juhász, S., Künzel, J., Spies, J., Shibata, A., Jeggo, P.A., and Löbrich, M. (2017). DNA double-strand break resection occurs during non-homologous end joining in G1 but is distinct from resection during homologous recombination. *Mol. Cell* 65, 671-684. e675.

Bohrer, S., Ades, L., Tajeddine, N., Hofmann, W., Kriener, S., Bug, G., Ottmann, O., Ruthardt, M., Galluzzi, L., and Fouassier, C. (2009). Suppression of the DNA damage response in acute myeloid leukemia versus myelodysplastic syndrome. *Oncogene* 28, 2205-2218.

Bradford, M., (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* 72. 248–254

Brosh Jr, R.M., and Bohr, V.A. (2007). Human premature aging, DNA repair and RecQ helicases. *Nucleic Acids Res.* 35, 7527-7544.

Bunting, S.F., Callén, E., Wong, N., Chen, H.-T., Polato, F., Gunn, A., Bothmer, A., Feldhahn, N., Fernandez-Capetillo, O., and Cao, L. (2010). 53BP1 inhibits homologous recombination in Brca1-deficient cells by blocking resection of DNA breaks. *Cell* 141, 243-254.

Byrd, A.K., and Raney, K.D. (2012). Superfamily 2 helicases. *Front. Biosci.* 17, 2070.

Cannan, W.J., and Pederson, D.S. (2016). Mechanisms and consequences of double-strand DNA break formation in chromatin. *J. Cell. Physiol.* 231, 3-14.

Cardoso, S.R., Ryan, G., Walne, A.J., Ellison, A., Lowe, R., Tummala, H., Rio-Machin, A., Collopy, L., Al Seraihi, A., and Wallis, Y. (2016). Germline heterozygous DDX41 variants in a subset of familial myelodysplasia and acute myeloid leukemia. *Leukemia* 30, 2083-2086.

Caruthers, J.M., and McKay, D.B. (2002). Helicase structure and mechanism. *Curr. Opin. Struct. Biol.* 12, 123-133.

Cavelier, C., Didier, C., Prade, N., Mansat-De Mas, V., Manenti, S., Recher, C., Demur, C., and Ducommun, B. (2009). Constitutive activation of the DNA damage signaling pathway in acute myeloid leukemia with complex karyotype: potential importance for checkpoint targeting therapy. *Cancer Res.* 69, 8652-8661.

Cerritelli, S. M., and Crouch, R. J. (2009). Ribonuclease H: the enzymes in eukaryotes. *FEBS J* 276(6), 1494-1505.

Chakraborty, P., Huang, J.T., and Hiom, K. (2018a). DHX9 helicase promotes R-loop formation in cells with impaired RNA splicing. *Nat. Commun.* 9, 1-14.

Chan, C.H., Chen, C.M., Lee, Y.W., and You, L.R. (2019). DNA Damage, Liver Injury, and Tumorigenesis: Consequences of DDX3X Loss. *Mol. Cancer Res.* 17, 555-566.

Chapman, J.R., Barral, P., Vannier, J.-B., Borel, V., Steger, M., Tomas-Loba, A., Sartori, A.A., Adams, I.R., Batista, F.D., and Boulton, S.J. (2013). RIF1 is essential for 53BP1-dependent nonhomologous end joining and suppression of DNA double-strand break resection. *Mol. Cell* 49, 858-871.

Chatterjee, N., and Walker, G.C. (2017). Mechanisms of DNA damage, repair, and mutagenesis. *Environ. Mol. Mutagen* 58, 235-263.

Chen, J., Ghorai, M.K., Kenney, G., and Stubbe, J. (2008). Mechanistic studies on bleomycin-mediated DNA damage: multiple binding modes can result in double-stranded DNA cleavage. *Nucleic Acids Res.* 36, 3781-3790.

Chen, L., Chen, J.-Y., Huang, Y.-J., Gu, Y., Qiu, J., Qian, H., Shao, C., Zhang, X., Hu, J., and Li, H. (2018). The augmented R-loop is a unifying mechanism for myelodysplastic syndromes induced by high-risk splicing factor mutations. *Mol. Cell* 69, 412-425. e416.

Chuang, R.-Y., Weaver, P.L., Liu, Z., and Chang, T.-H. (1997). Requirement of the DEAD-Box protein ded1p for messenger RNA translation. *Science* 275, 1468-1471.

Ciccia, A., and Elledge, S.J. (2010). The DNA damage response: making it safe to play with knives. *Mol. Cell* 40, 179-204.

Cristini, A., Groh, M., Kristiansen, M.S., and Gromak, N. (2018). RNA/DNA Hybrid Interactome Identifies DXH9 as a Molecular Player in Transcriptional Termination and R-Loop-Associated DNA Damage. *Cell Rep.* 23, 1891-1905.

Cristini, A., Ricci, G., Britton, S., Salimbeni, S., Huang, S.-y.N., Marinello, J., Calsou, P., Pommier, Y., Favre, G., and Capranico, G. (2019). Dual processing of R-Loops and topoisomerase I induces transcription-dependent DNA double-strand breaks. *Cell Rep.* 28, 3167-3181. e3166.

Dillingham, M.S. (2011). Superfamily I helicases as modular components of DNA-processing machines. *Biochem. Soc. Trans.* 39, 413-423.

Ding, L., Ley, T.J., Larson, D.E., Miller, C.A., Koboldt, D.C., Welch, J.S., Ritchey, J.K., Young, M.A., Lamprecht, T., and McLellan, M.D. (2012). Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing. *Nature* 481, 506-510.

Domínguez-Sánchez, M.S., Barroso, S., Gómez-González, B., Luna, R., and Aguilera, A. (2011). Genome instability and transcription elongation impairment in human cells depleted of THO/TREX. *PLoS Genet.* 7, e1002386.

Duployez, N., Marceau-Renaut, A., Boissel, N., Petit, A., Bucci, M., Geffroy, S., Lapillonne, H., Renneville, A., Ragu, C., and Figeac, M. (2016). Comprehensive mutational profiling of core binding factor acute myeloid leukemia. *Blood.* 127, 2451-2459.

Dutta, D., Shatalin, K., Epshtein, V., Gottesman, M. E., & Nudler, E. (2011). Linking RNA polymerase backtracking to genome instability in *E. coli*. *Cell* 146(4), 533-543.

Fenwarth, L., Caulier, A., Lachaier, E., Goursaud, L., Marceau-Renaut, A., Fournier, E., Lebon, D., Boyer, T., Berthon, C., and Marolleau, J.-P. (2021). Hereditary Predisposition to Acute Myeloid Leukemia in Older Adults. *HemaSphere* 5, e552.

Fuller-Pace, F.V. (2013). DEAD box RNA helicase functions in cancer. *RNA Biol.* 10, 121-132.

Gaillard, H., García-Muse, T., and Aguilera, A. (2015). Replication stress and cancer. *Nat. Rev. Cancer* 15, 276-289.

García-Rubio, M., Aguilera, P., Lafuente-Barquero, J., Ruiz, J.F., Simon, M.-N., Geli, V., Rondón, A.G., and Aguilera, A. (2018). Yra1-bound RNA–DNA hybrids cause orientation-independent transcription–replication collisions and telomere instability. *Genes Dev.* 32, 965-977.

Gatfield, D., Le Hir, H., Schmitt, C., Braun, I.C., Köcher, T., Wilm, M., and Izaurralde, E. (2001). The DExH/D box protein HEL/UAP56 is essential for mRNA nuclear export in *Drosophila*. *Curr. Biol.* 11, 1716-1721.

Ginno, P.A., Lott, P.L., Christensen, H.C., Korf, I., and Chédin, F. (2012). R-loop formation is a distinctive characteristic of unmethylated human CpG island promoters. *Mol. Cell* 45, 814-825.

Gorbalenya, A.E., and Koonin, E.V. (1993). Helicases: amino acid sequence comparisons and

structure-function relationships. *Curr. Opin. Struct.* 3, 419-429.

Grosjean-Raillard, J., Tailler, M., Ades, L., Perfettini, J., Fabre, C., Braun, T., De Botton, S., Fenaux, P., and Kroemer, G. (2009). ATM mediates constitutive NF- κ B activation in high-risk myelodysplastic syndrome and acute myeloid leukemia. *Oncogene* 28, 1099-1109.

Gunn, A., and Stark, J.M. (2012). I-SceI-based assays to examine distinct repair outcomes of mammalian chromosomal double strand breaks. *Methods Mol. Biol.* 920, 379-391.

Guo, M., Hundseth, K., Ding, H., Vidhyasagar, V., Inoue, A., Nguyen, C.H., Zain, R., Lee, J.S., and Wu, Y. (2015). A Distinct Triplex DNA Unwinding Activity of ChlR1 Helicase. *J. Biol. Chem.* 290, 5174-5189.

Haeusler, A.R., Donnelly, C.J., Periz, G., Simko, E.A., Shaw, P.G., Kim, M.-S., Maragakis, N.J., Troncoso, J.C., Pandey, A., and Sattler, R. (2014). C9orf72 nucleotide repeat structures initiate molecular cascades of disease. *Nature* 507, 195-200.

Haferlach, C., Dicker, F., Herholz, H., Schnittger, S., Kern, W., and Haferlach, T. (2008). Mutations of the TP53 gene in acute myeloid leukemia are strongly associated with a complex aberrant karyotype. *Leukemia* 22, 1539-1541.

Halazonetis, T.D., Gorgoulis, V.G., and Bartek, J. (2008). An Oncogene -induced DNA damage model for cancer development. *Science* 319, 1352-1355.

Hamperl, S., and Cimprich, K.A. (2014). The contribution of co-transcriptional RNA: DNA hybrid structures to DNA damage and genome instability. *DNA Repair* 19, 84-94.

Hegazy, Y.A., Fernando, C.M., and Tran, E.J. (2020). The balancing act of R-loop biology: The good, the bad, and the ugly. *J. Biol. Chem.* 295, 905-913.

Hodroj, D., Recolin, B., Serhal, K., Martinez, S., Tsanov, N., Abou Merhi, R., and Maiorano, D. (2017). An ATR-dependent function for the Ddx19 RNA helicase in nuclear R-loop metabolism. *EMBO J.* 36, 1182-1198.

Hoeijmakers, J.H. (2001). Genome maintenance mechanisms for preventing cancer. *Nature* 411, 366-374.

Horibe, S., Takagi, M., Unno, J., Nagasawa, M., Morio, T., Arai, A., Miura, O., Ohta, M., Kitagawa, M., and Mizutani, S. (2007). DNA damage check points prevent leukemic transformation in myelodysplastic syndrome. *Leukemia* 21, 2195-2198.

Hosono, N. (2019). Genetic defects of chromosome 5q and 7q in myeloid neoplasms. *Rinsho ketsueki.* 60, 800-809.

Iacobucci, I., Wen, J., Meggendorfer, M., Choi, J.K., Shi, L., Pounds, S.B., Carmichael, C.L., Masih, K.E., Morris, S.M., and Lindsley, R.C. (2019). Genomic subtyping and therapeutic

targeting of acute erythroleukemia. *Nat. Genet.* 51, 694-704.

Ishikawa, H., and Barber, G.N. (2008). STING is an endoplasmic reticulum adaptor that facilitates innate immune signalling. *Nature* 455, 674-678.

Jackson, R.N., Lavin, M., Carter, J., and Wiedenheft, B. (2014). Fitting CRISPR-associated Cas3 into the helicase family tree. *Curr. Opin. Struct.* 24, 106-114.

Jahn, A., Rane, G., Paszkowski-Rogacz, M., Sayols, S., Bluhm, A., Han, C.-T., Draškovič, I., Londoño, A., Kumar, A.P., and Buchholz, F. (2018). ZBTB48 is both a vertebrate telomere-binding protein and a transcriptional activator. *EMBO Rep.* 18, 929-946.

Jankowsky, E. (2011). RNA helicases at work: binding and rearranging. *Trends Biochem. Sci.* 36, 19-29.

Jiang, Y., Dunbar, A., Gondek, L.P., Mohan, S., Rataul, M., O'Keefe, C., Sekeres, M., Saunthararajah, Y., and Maciejewski, J.P. (2009). Aberrant DNA methylation is a dominant mechanism in MDS progression to AML. *Blood* 113, 1315-1325.

Jiang, Y., Zhu, Y., Liu, Z.J., and Ouyang, S. (2017). The emerging roles of the DDX41 protein in immunity and diseases. *Protein Cell* 8, 83-89.

Jiricny, J. (2006). The multifaceted mismatch-repair system. *Nat. Rev. Mol. Cell. Biol* 7, 335-346.

Joshi, D., Korgaonkar, S., Shanmukhaiah, C., and Vundinti, B.R. (2016). Association of XPD (Lys751Gln) and XRCC1 (Arg280His) gene polymorphisms in myelodysplastic syndrome. *Ann. Hematol.* 95, 79-85.

Jurica, M.S., Licklider, L.J., Gygi, S.P., Grigorieff, N., and Moore, M.J. (2002). Purification and characterization of native spliceosomes suitable for three-dimensional structural analysis. *RNA* 8, 426-439.

Kabeche, L., Nguyen, H.D., Buisson, R., and Zou, L. (2018). A mitosis-specific and R loop-driven ATR pathway promotes faithful chromosome segregation. *Science* 359, 108-114.

Kadono, M., Kanai, A., Nagamachi, A., Shinriki, S., Kawata, J., Iwato, K., Kyo, T., Oshima, K., Yokoyama, A., and Kawamura, T. (2016). Biological implications of somatic DDX41 p. R525H mutation in acute myeloid leukemia. *Exp. Hematol.* 44, 745-754. e744.

Kahlina, K., Goren, I., Pfeilschifter, J., and Frank, S. (2004). p68 DEAD box RNA helicase expression in keratinocytes. Regulation, nucleolar localization, and functional connection to proliferation and vascular endothelial growth factor gene expression. *J. Biol. Chem.* 279, 44872-44882.

Kato, R., Miyagawa, K., and Yasuhara, T. (2019). The role of R-loops in transcription-associated DNA double-strand break repair. *Mol. Cellular Oncol.* 6, 1542244.

Kefala, M., Papageorgiou, S.G., Kontos, C.K., Economopoulou, P., Tsanas, A., Pappa, V., Panayiotides, I.G., Gorgoulis, V.G., Patsouris, E., and Foukas, P.G. (2013). Increased expression of phosphorylated NBS1, a key molecule of the DNA damage response machinery, is an adverse prognostic factor in patients with de novo myelodysplastic syndromes. *Leukemia Res.* 37, 1576-1582.

Kemp, M.G., and Sancar, A. (2012). DNA excision repair: where do all the dimers go? *Cell Cycle* 11, 2997-3002.

Kim, N., and Jinks-Robertson, S. (2012). Transcription as a source of genome instability. *Nat. Rev. Genet.* 13, 204-214.

Kim, S., Kang, N., Park, S.H., Wells, J., Hwang, T., Ryu, E., Kim, B.G., Hwang, S., Kim, S.J., Kang, S., *et al.* (2020). ATAD5 restricts R-loop formation through PCNA unloading and RNA helicase maintenance at the replication fork. *Nucleic Acids Res.* 48, 7218-7238.

Koo, C.X.e., Kobiyama, K., Shen, Y.J., LeBert, N., Ahmad, S., Khatoo, M., Aoshi, T., Gasser, S., and Ishii, K.J. (2015). RNA polymerase III regulates cytosolic RNA: DNA hybrids and intracellular microRNA expression. *J. Biol. Chem.* 290, 7463-7473.

Kouyama, Y., Masuda, T., Fujii, A., Ogawa, Y., Sato, K., Tobo, T., Wakiyama, H., Yoshikawa, Y., Noda, M., and Tsuruda, Y. (2019). Oncogenic splicing abnormalities induced by DEAD-Box Helicase 56 amplification in colorectal cancer. *Cancer Sci.* 110, 3132-3144.

Lang, K.S., Hall, A.N., Merrih, C.N., Ragheb, M., Tabakh, H., Pollock, A.J., Woodward, J.J., Dreifus, J.E., and Merrih, H. (2017). Replication-transcription conflicts generate R-loops that orchestrate bacterial stress survival and pathogenesis. *Cell* 170, 787-799.

Lasko, P. (2013). The DEAD-box helicase Vasa: evidence for a multiplicity of functions in RNA processes and developmental biology. *Biochim Biophys Acta Gene Regul. Mech.* 1829, 810-816.

Lee, K.-G., Kim, S.S.-Y., Kui, L., Voon, D.C.-C., Mauduit, M., Bist, P., Bi, X., Pereira, N.A., Liu, C., and Sukumaran, B. (2015a). Bruton's tyrosine kinase phosphorylates DDX41 and activates its binding of dsDNA and STING to initiate type 1 interferon response. *Cell Rep.* 10, 1055-1065.

Lewinsohn, M., Brown, A.L., Weinel, L.M., Phung, C., Rafidi, G., Lee, M.K., Schreiber, A.W., Feng, J., Babic, M., and Chong, C.-E. (2016). Novel germ line DDX41 mutations define families with a lower age of MDS AND AML onset and lymphoid malignancies. *Blood* 127, 1017-1023.

Li, L., Germain, D.R., Poon, H.Y., Hildebrandt, M.R., Monckton, E.A., McDonald, D., Hendzel, M.J., and Godbout, R. (2016). DEAD Box 1 Facilitates Removal of RNA and Homologous Recombination at DNA Double-Strand Breaks. *Mol. Cell Biol.* 36, 2794-2810.

Li, L., Monckton, E.A., and Godbout, R. (2008). A role for DEAD box 1 at DNA double-strand

breaks. *Mol. Cell Biol.* 28, 6413-6425.

Li, R., Sobreira, N., Witmer, P.D., Pratz, K.W., and Braunstein, E.M. (2016). Two novel germline DDX41 mutations in a family with inherited myelodysplasia/acute myeloid leukemia. *Haematologica* 101, e228.

Lindahl, T., and Wood, R.D. (1999). Quality control by DNA repair. *Science* 286, 1897-1905.

Linder, P. (2006). Dead-box proteins: a family affair—active and passive players in RNP-remodeling. *Nucleic Acids Res.* 34, 4168-4180.

Linder, P., and Fuller-Pace, F.V. (2013). Looking back on the birth of DEAD-box RNA helicases. *Biochim Biophys Acta Gene Regul. Mech.* 1829, 750-755.

Liu, H.-Y., Nefsky, B.S., and Walworth, N.C. (2002). The Ded1 DEAD box helicase interacts with Chk1 and Cdc2. *J. Biol. Chem.* 277, 2637-2643.

Lohman, T.M., Tomko, E.J., and Wu, C.G. (2008). Non-hexameric DNA helicases and translocases: mechanisms and regulation. *Nat Rev Mol Cell Biol.* 9, 391-401.

Loomis, E.W., Sanz, L.A., Chédin, F., and Hagerman, P.J. (2014). Transcription-associated R-loop formation across the human FMR1 CGG-repeat region. *PLoS Genet.* 10, e1004294.

Lu, W.-T., Hawley, B.R., Skalka, G.L., Baldock, R.A., Smith, E.M., Bader, A.S., Malewicz, M., Watts, F.Z., Wilczynska, A., and Bushell, M. (2018). Drosha drives the formation of DNA: RNA hybrids around DNA break sites to facilitate DNA repair. *Nat. Commun.* 9, 1-13.

Maciejewski, J.P., Padgett, R.A., Brown, A.L., and Müller-Tidow, C. (2017). DDX41-related myeloid neoplasia. *Semin. Hematol.*, 54, 94-97.

Magnusdottir, A., Johansson, I., Dahlgren, L.-G., Nordlund, P., and Berglund, H. (2009). Enabling IMAC purification of low abundance recombinant proteins from *E. coli* lysates. *Nat. Methods* 6, 477-478.

Malachin, G., Reiten, M.R., Salvesen, Ø., Aanes, H., Kamstra, J.H., Skovgaard, K., Heegaard, P.M., Ersdal, C., Espenes, A., and Tranulis, M.A. (2017). Loss of prion protein induces a primed state of type I interferon-responsive genes. *PLoS One* 12, e0179881.

Maréchal, A., and Zou, L. (2013). DNA damage sensing by the ATM and ATR kinases. *Cold Spring Harb. Perspect. Biol.* 5, a012716.

Meyer, T., Jahn, N., Lindner, S., Röhner, L., Dolnik, A., Weber, D., Scheffold, A., Köpff, S., Paschka, P., and Gaidzik, V.I. (2020). Functional characterization of BRCC3 mutations in acute myeloid leukemia with t (8; 21)(q22; q22. 1). *Leukemia* 34, 404-415.

Milek, M., Imami, K., Mukherjee, N., Bortoli, F., Zinnall, U., Hazapis, O., Trahan, C., Oeffinger,

M., Heyd, F., Ohler, U., *et al.* (2017). DDX54 regulates transcriptome dynamics during DNA damage response. *Genome Res.* 27, 1344-1359.

Mischo, H.E., Gómez-González, B., Grzechnik, P., Rondón, A.G., Wei, W., Steinmetz, L., Aguilera, A., and Proudfoot, N.J. (2011). Yeast Sen1 helicase protects the genome from transcription-associated instability. *Mol. Cell* 41, 21-32.

Mischo, H.E., Hemmerich, P., Grosse, F., and Zhang, S. (2005). Actinomycin D induces histone gamma-H2AX foci and complex formation of gamma-H2AX with Ku70 and nuclear DNA helicase II. *J. Biol. Chem.* 280, 9586-9594.

Missel, A., Souza, A.E., Nörskau, G., and Göringer, H. (1997). Disruption of a gene encoding a novel mitochondrial DEAD-box protein in *Trypanosoma brucei* affects edited mRNAs. *Mol. Cell Biol.* 17, 4895-4903.

Mosler, T., Conte, F., Mikicic, I., Kreim, N., Möckel, M., Flach, J., Luke, B., and Beli, P. (2021). R-loop proximity proteomics identifies a role of DDX41 in transcription-associated genomic instability. *Nat. Portfolio*. doi:10.21203/rs.3.rs-337351/v1.

Mossner, M., Jann, J.-C., Wittig, J., Nolte, F., Fey, S., Nowak, V., Obländer, J., Pressler, J., Palme, I., and Xanthopoulos, C. (2016). Mutational hierarchies in myelodysplastic syndromes dynamically adapt and evolve upon therapy response and failure. *Blood* 128, 1246-1259.

Namiki, Y., and Zou, L. (2006). ATRIP associates with replication protein A-coated ssDNA through multiple interactions. *Proc. Natl. Acad. Sci. U.S.A.* 103, 580-585.

Ohle, C., Tesorero, R., Schermann, G., Dobrev, N., Sinning, I., and Fischer, T. (2016). Transient RNA-DNA hybrids are required for efficient double-strand break repair. *Cell* 167, 1001-1013. e1007.

Okamoto, Y., Abe, M., Itaya, A., Tomida, J., Ishiai, M., Takaori-Kondo, A., Taoka, M., Isobe, T., and Takata, M. (2019). FANCD2 protects genome stability by recruiting RNA processing enzymes to resolve R-loops during mild replication stress. *FEBS J.* 286, 139-150.

Omura, H., Oikawa, D., Nakane, T., Kato, M., Ishii, R., Ishitani, R., Tokunaga, F., and Nureki, O. (2016). Structural and Functional Analysis of DDX41: a bispecific immune receptor for DNA and cyclic dinucleotide. *Sci. Rep.* 6.

Papaemmanuil, E., Gerstung, M., Malcovati, L., Tauro, S., Gundem, G., Van Loo, P., Yoon, C.J., Ellis, P., Wedge, D.C., and Pellagatti, A. (2013). Clinical and biological implications of driver mutations in myelodysplastic syndromes. *Blood* 122, 3616-3627.

Parsyan, A., Svitkin, Y., Shahbazian, D., Gkogkas, C., Lasko, P., Merrick, W.C., and Sonenberg, N. (2011). mRNA helicases: the tacticians of translational control. *Nat. Rev. Mol. Cell. Biol.* 12, 235-245.

Parvatiyar, K., Zhang, Z., Teles, R.M., Ouyang, S., Jiang, Y., Iyer, S.S., Zaver, S.A., Schenk, M., Zeng, S., and Zhong, W. (2012). The helicase DDX41 recognizes the bacterial secondary messengers cyclic di-GMP and cyclic di-AMP to activate a type I interferon immune response. *Nat. Immunol.* *13*, 1155-1161.

Patel, S.S., and Picha, K.M. (2000). Structure and function of hexameric helicases. *Annu. Rev. Biochem.* *69*, 651-697.

Pause, A., Methot, N., Svitkin, Y., Merrick, W., and Sonenberg, N. (1994). Dominant negative mutants of mammalian translation initiation factor eIF-4A define a critical role for eIF-4F in cap-dependent and cap-independent initiation of translation. *EMBO J.* *13*, 1205-1215.

Pause, A., and Sonenberg, N. (1992). Mutational analysis of a DEAD box RNA helicase: the mammalian translation initiation factor eIF-4A. *EMBO J.* *11*, 2643-2654.

Pérez-Calero, C., Bayona-Feliu, A., Xue, X., Barroso, S.I., Muñoz, S., González-Basallote, V.M., Sung, P., and Aguilera, A. (2020). UAP56/DDX39B is a major cotranscriptional RNA–DNA helicase that unwinds harmful R loops genome-wide. *Genes Dev.* *34*, 898-912.

Peters, D., Radine, C., Reese, A., Budach, W., Sohn, D., and Jänicke, R.U. (2017). The DEAD-box RNA helicase DDX41 is a novel repressor of p21WAF1/CIP1 mRNA translation. *J. Biol. Chem.* *292*, 8331-8341.

Polprasert, C., Schulze, I., Sekeres, M.A., Makishima, H., Przychodzen, B., Hosono, N., Singh, J., Padgett, R.A., Gu, X., and Phillips, J.G. (2015). Inherited and somatic defects in DDX41 in myeloid neoplasms. *Cancer Cell* *27*, 658-670.

Popp, H.D., Naumann, N., Brendel, S., Henzler, T., Weiss, C., Hofmann, W.-K., and Fabarius, A. (2017). Increase of DNA damage and alteration of the DNA damage response in myelodysplastic syndromes and acute myeloid leukemias. *Leukemia Res.* *57*, 112-118.

Py, B., Higgins, C.F., Krisch, H.M., and Carpousis, A.J. (1996). A DEAD-box RNA helicase in the *Escherichia coli* RNA degradosome. *Nature* *381*, 169-172.

Qin, K., Jian, D., Xue, Y., Cheng, Y., Zhang, P., Wei, Y., Zhang, J., Xiong, H., Zhang, Y., and Yuan, X. (2021). DDX41 regulates the expression and alternative splicing of genes involved in tumorigenesis and immune response. *Oncol. Rep.* *45*, 1213-1225.

Ran, F. A., Hsu, P. D., Wright, J., Agarwala, V., Scott, D. A., & Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. *Nature protocols* *8(11)*, 2281-2308.

Ranoa, D.R.E., Parekh, A.D., Pitroda, S.P., Huang, X., Darga, T., Wong, A.C., Huang, L., Andrade, J., Staley, J.P., and Satoh, T. (2016). Cancer therapies activate RIG-I-like receptor pathway through endogenous non-coding RNAs. *Oncotarget* *7*, 26496.

Redon, C.E., Dickey, J.S., Bonner, W.M., and Sedelnikova, O.A. (2009). γ -H2AX as a

biomarker of DNA damage induced by ionizing radiation in human peripheral blood lymphocytes and artificial skin. *Adv. Space Res.* 43, 1171-1178.

Ria, R., Moschetta, M., Reale, A., Mangialardi, G., Castrovilli, A., Vacca, A., and Dammacco, F. (2009). Managing myelodysplastic symptoms in elderly patients. *Clin. Interv. Aging* 4, 413.

Ryan, E.L., Hollingworth, R., and Grand, R.J. (2016). Activation of the DNA damage response by RNA viruses. *Biomolecules* 6, 2.

Rothkamm, K., Barnard, S., Moquet, J., Ellender, M., Rana, Z., and Burdak-Rothkamm, S. (2015). DNA damage foci: Meaning and significance. *Environ. Mol. Mutagen.* 56(6), 491-504.

Sanchez, A., de Vivo, A., Tonzi, P., Kim, J., Huang, T.T., and Kee, Y. (2020). Transcription-replication conflicts as a source of common fragile site instability caused by BMI1-RNF2 deficiency. *PLoS Genet.* 16, e1008524.

Sanjana, N. E., Shalem, O., & Zhang, F. (2014). Improved vectors and genome-wide libraries for CRISPR screening. *Nature methods* 11(8), 783-784.

Santos-Pereira, J.M., and Aguilera, A. (2015). R loops: new modulators of genome dynamics and function. *Nat. Rev. Genet.* 16, 583-597.

Satoh, Y., Matsumura, I., Tanaka, H., Harada, H., Harada, Y., Matsui, K., Shibata, M., Mizuki, M., and Kanakura, Y. (2012). C-terminal mutation of RUNX1 attenuates the DNA-damage repair response in hematopoietic stem cells. *Leukemia* 26, 303-311.

Saygin, C., and Godley, L.A. (2021). Genetics of Myelodysplastic Syndromes. *Cancers* 13, 3380.

Schwab, R.A., Nieminuszczy, J., Shah, F., Langton, J., Martinez, D.L., Liang, C.-C., Cohn, M.A., Gibbons, R.J., Deans, A.J., and Niedzwiedz, W. (2015). The Fanconi anemia pathway maintains genome stability by coordinating replication and transcription. *Mol. Cell* 60, 351-361.

Shah, N., Inoue, A., Woo, L.S., Beishline, K., Lahti, J.M., and Noguchi, E. (2013). Roles of ChlR1 DNA helicase in replication recovery from DNA damage. *Exp. Cell Res.* 319, 2244-2253.

Singleton, M.R., Dillingham, M.S., and Wigley, D.B. (2007). Structure and mechanism of helicases and nucleic acid translocases. *Annu. Rev. Biochem.* 76, 23-50.

Skourti-Stathaki, K., Proudfoot, N.J., and Gromak, N. (2011). Human senataxin resolves RNA/DNA hybrids formed at transcriptional pause sites to promote Xrn2-dependent termination. *Mol. Cell* 42, 794-805.

Smolka, J.A., Sanz, L.A., Hartono, S.R., and Chédin, F. (2021). Recognition of RNA by the S9.6 antibody creates pervasive artifacts when imaging RNA: DNA hybrids. *J. Cell Biol.* 220, e202004079.

- Sollier, J., and Cimprich, K.A. (2015). Breaking bad: R-loops and genome integrity. *Trends Cell Biol.* *25*, 514-522.
- Sollier, J., Stork, C.T., García-Rubio, M.L., Paulsen, R.D., Aguilera, A., and Cimprich, K.A. (2014). Transcription-coupled nucleotide excision repair factors promote R-loop-induced genome instability. *Mol. Cell* *56*, 777-785.
- Song, C., Hotz-Wagenblatt, A., Voit, R., and Grummt, I. (2017). SIRT7 and the DEAD-box helicase DDX21 cooperate to resolve genomic R loops and safeguard genome stability. *Genes Dev.* *10*.
- Sridhara, S.C., Carvalho, S., Grosso, A.R., Gallego-Paez, L.M., Carmo-Fonseca, M., and de Almeida, S.F. (2017). Transcription dynamics prevent RNA-mediated genomic instability through SRPK2-dependent DDX23 phosphorylation. *Cell Rep.* *18*, 334-343.
- Sun, M., Zhou, T., Jonasch, E., and Jope, R.S. (2013). DDX3 regulates DNA damage-induced apoptosis and p53 stabilization. *Biochim. Biophys. Acta Mol. Cell Res.* *1833*, 1489-1497.
- Szappanos, D., Tschismarov, R., Perlot, T., Westermayer, S., Fischer, K., Platanitis, E., Kallinger, F., Novatchkova, M., Lassnig, C., and Müller, M. (2018). The RNA helicase DDX3X is an essential mediator of innate antimicrobial immunity. *PLoS Pathog.* *14*, e1007397.
- Takaoka, K., Kawazu, M., Koya, J., Yoshimi, A., Masamoto, Y., Maki, H., Toya, T., Kobayashi, T., Nannya, Y., and Arai, S. (2019). A germline HLTF mutation in familial MDS induces DNA damage accumulation through impaired PCNA polyubiquitination. *Leukemia* *33*, 1773-1782.
- Tanaka, Y., and Chen, Z.J. (2012). STING specifies IRF3 phosphorylation by TBK1 in the cytosolic DNA signaling pathway. *Sci. Signal.* *5*, ra20-ra20.
- Tawana, K., Drazer, M.W., and Churpek, J.E. (2018). Universal genetic testing for inherited susceptibility in children and adults with myelodysplastic syndrome and acute myeloid leukemia: are I there yet? *Leukemia* *32*, 1482-1492.
- Tawana, K., and Fitzgibbon, J. (2016). Inherited DDX41 mutations: 11 genes and counting. *Blood* *127*, 960-961.
- Tubbs, A., and Nussenzweig, A. (2017). Endogenous DNA damage as a source of genomic instability in cancer. *Cell* *168*, 644-656.
- Turnell, A.S., and Grand, R.J. (2012). DNA viruses and the cellular DNA-damage response. *J. Gen. Virol.* *93*, 2076-2097.
- Quesada, A. E., Routbort, M. J., DiNardo, C. D., Bueso-Ramos, C. E., Kanagal-Shamanna, R., Khoury, J. D., ... & Patel, K. P. (2019). DDX41 mutations in myeloid neoplasms are associated with male gender, TP53 mutations and high-risk disease. *Am. J. Hematol.* *94*(7), 757-766.

- Valentini, M., and Linder, P. (2021). Happy Birthday: 30 Years of RNA Helicases. *Methods Mol. Biol.* 2209, 17-34.
- Venema, J., and Tollervey, D. (1999). Ribosome synthesis in *Saccharomyces cerevisiae*. *Annu. Rev. Genet.* 33, 261-311.
- Weinreb, J.T., Ghazale, N., Pradhan, K., Gupta, V., Potts, K.S., Tricoli, B., Daniels, N.J., Padgett, R.A., De Oliveira, S., and Verma, A. (2021). Excessive R-loops trigger an inflammatory cascade leading to increased HSPC production. *Dev. Cell* 56, 627-640. e625.
- Weinreb, J.T., Gupta, V., Sharvit, E., Weil, R., and Bowman, T.V. (2021). Ddx41 inhibition of DNA damage signaling permits erythroid progenitor expansion in zebrafish. *Haematologica PMID: 33763998*.
- Yasuhara, T., Kato, R., Hagiwara, Y., Shiotani, B., Yamauchi, M., Nakada, S., Shibata, A., and Miyagawa, K. (2018). Human Rad52 promotes XPG-mediated R-loop processing to initiate transcription-associated homologous recombination repair. *Cell* 175, 558-570. e511.
- Yu, Z., Mersaoui, S.Y., Guitton-Sert, L., Coulombe, Y., Song, J., Masson, J.-Y., and Richard, S. (2020). DDX5 resolves R-loops at DNA double-strand breaks to promote DNA repair and avoid chromosomal deletions. *NAR Cancer* 2, zcaa028.
- Zhang, Z., Kim, T., Bao, M., Facchinetti, V., Jung, S.Y., Ghaffari, A.A., Qin, J., Cheng, G., and Liu, Y.-J. (2011a). DDX1, DDX21, and DHX36 helicases form a complex with the adaptor molecule TRIF to sense dsRNA in dendritic cells. *Immunity* 34, 866-878.
- Zhang, Z., Yuan, B., Bao, M., Lu, N., Kim, T., and Liu, Y.J. (2011c). The helicase DDX41 senses intracellular DNA mediated by the adaptor STING in dendritic cells. *Nat. Immunol.* 12, 959-965.
- Zhou, B.-B.S., and Elledge, S.J. (2000). The DNA damage response: putting checkpoints in perspective. *Nature* 408, 433-439.

Permission to use figures

For Figure 1

12/8/21, 10:51 AM

Mali - Aggarwal, Aanchal - Outlook


Thank you for your order with RightsLink / Elsevier

no-reply@copyright.com <no-reply@copyright.com>

Thu 11/18/2021 2:40 PM

To: Aggarwal, Aanchal <aanchal.aggarwal@usask.ca>

CAUTION: External to USask. Verify sender and use caution with links and attachments. Forward suspicious emails to phishing@usask.ca

 Header

Thank you for your order!

Dear Ms. Aanchal Aggarwal,

Thank you for placing your order through Copyright Clearance Center's RightsLink® service.

Order Summary

Licensee: University of Saskatchewan
Order Date: Nov 18, 2021
Order Number: 5192081113661
Publication: Current Opinion in Structural Biology
Title: Fitting CRISPR-associated Cas3 into the Helicase Family Tree
Type of Use: reuse in a thesis/dissertation
Order Total: 0.00 CAD

View or print complete [details](#) of your order and the publisher's terms and conditions.

Sincerely,

Copyright Clearance Center

For Figure 2

Thank you for your order with RightsLink / Springer Nature

no-reply@copyright.com <no-reply@copyright.com>

Thu 11/18/2021 2:46 PM

To: Aggarwal, Aanchal <aanchal.aggarwal@usask.ca>

CAUTION: External to USask. Verify sender and use caution with links and attachments. Forward suspicious emails to phishing@usask.ca



Thank you for your order!

Dear Ms. Aanchal Aggarwal,

Thank you for placing your order through Copyright Clearance Center's RightsLink® service.

Order Summary

Licensee:	University of Saskatchewan
Order Date:	Nov 18, 2021
Order Number:	5192081463124
Publication:	Nature Reviews Molecular Cell Biology
Title:	mRNA helicases: the tacticians of translational control
Type of Use:	Thesis/Dissertation
Order Total:	0.00 CAD

View or print complete [details](#) of your order and the publisher's terms and conditions.

Sincerely,

Copyright Clearance Center

For Figure 5

Thank you for your order with RightsLink / Springer Nature

no-reply@copyright.com <no-reply@copyright.com>

Thu 11/18/2021 2:53 PM

To: Aggarwal, Aanchal <aanchal.aggarwal@usask.ca>

CAUTION: External to USask. Verify sender and use caution with links and attachments. Forward suspicious emails to phishing@usask.ca

 Header

Thank you for your order!

Dear Ms. Aanchal Aggarwal,

Thank you for placing your order through Copyright Clearance Center's RightsLink® service.

Order Summary

Licensee: University of Saskatchewan
Order Date: Nov 18, 2021
Order Number: 5192090415790
Publication: Nature
Title: Genome maintenance mechanisms for preventing cancer
Type of Use: Thesis/Dissertation
Order Total: 0.00 CAD

View or print complete [details](#) of your order and the publisher's terms and conditions.

Sincerely,

Copyright Clearance Center

For Figure 6

Thank You for Your Order on Marketplace™

no-reply@copyright.com <no-reply@copyright.com>

Thu 11/18/2021 3:20 PM

To: Aggarwal, Aanchal <aanchal.aggarwal@usask.ca>

CAUTION: External to USask. Verify sender and use caution with links and attachments. Forward suspicious emails to phishing@usask.ca

 [Market place logo](#)

Dear Aanchal Aggarwal,

Thank you for placing your order on [Marketplace™](#).

Order Summary:

Order date: 18 Nov 2021

Order number: 1162711

No. of items: 1

Order total: 0.00 CAD

Billing Summary:

Payment method: Invoice

An invoice will be generated and emailed within 24 hours.

To view your order details, click the following link, sign in, and search for your order:

[Manage Account](#).

How was your experience? [Click here to give us feedback](#)

Please do not reply to this message.

To speak with a Customer Service Representative, call +1-855-239-3415 toll free or +1-978-646-2600 (24 hours a day), or email your questions and comments to support@copyright.com.

Sincerely,

The CCC Marketplace Team

For Figure 7

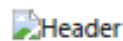
Thank you for your order with RightsLink / Elsevier

no-reply@copyright.com <no-reply@copyright.com>

Wed 12/8/2021 11:33 AM

To: Aggarwal, Aanchal <aanchal.aggarwal@usask.ca>

CAUTION: External to USask. Verify sender and use caution with links and attachments. Forward suspicious emails to phishing@usask.ca



Thank you for your order!

Dear Ms. Aanchal Aggarwal,

Thank you for placing your order through Copyright Clearance Center's RightsLink® service.

Order Summary

Licensee: University of Saskatchewan
Order Date: Dec 8, 2021
Order Number: 5204291332068
Publication: DNA Repair
Title: The contribution of co-transcriptional RNA:DNA hybrid structures to DNA damage and genome instability
Type of Use: reuse in a thesis/dissertation
Order Total: 0.00 CAD

View or print complete [details](#) of your order and the publisher's terms and conditions.

Sincerely,

Copyright Clearance Center

For Figure 8



[Genes Dev.](#) 2018 Jul 1; 32(13-14): 965–977.

PMCID: PMC6075034

doi: [10.1101/gad.311274.117](https://doi.org/10.1101/gad.311274.117)

PMID: [29954833](https://pubmed.ncbi.nlm.nih.gov/29954833/)

Yra1-bound RNA–DNA hybrids cause orientation-independent transcription–replication collisions and telomere instability

[María García-Rubio](#),^{1,3} [Paula Aguilera](#),^{2,3} [Juan Lafuente-Barquero](#),¹ [José F. Ruiz](#),¹ [Marie-Noelle Simon](#),² [Vincent Geli](#),² [Ana G. Rondón](#),¹ and [Andrés Aguilera](#)¹

► [Author information](#) ► [Article notes](#) ► [Copyright and License information](#) [Disclaimer](#)

[Copyright](#) © 2018 García-Rubio et al.; Published by Cold Spring Harbor Laboratory Press

This article is distributed exclusively by Cold Spring Harbor Laboratory Press for the first six months after the full-issue publication date (see <http://genesdev.cshlp.org/site/misc/terms.xhtml>). After six months, it is available under a Creative Commons License (Attribution-NonCommercial 4.0 International), as described at

<http://creativecommons.org/licenses/by-nc/4.0/>



[Share your work](#)

[Use & remix](#)

[What We do](#)

[Blog](#)



Attribution-NonCommercial 4.0 International (CC BY-NC 4.0)

This is a human-readable summary of (and not a substitute for) the [license](#). [Disclaimer](#).

You are free to:

Share — copy and redistribute the material in any medium or format

Adapt — remix, transform, and build upon the material

The licensor cannot revoke these freedoms as long as you follow the license terms.

For Figure 9



Published by the American Society for
Biochemistry and Molecular Biology

[J Biol Chem](#). 2020 Jan 24; 295(4): 905–913.

PMCID: PMC6983857

Published online 2019 Dec 16. doi: [10.1074/jbc.REV119.011353](https://doi.org/10.1074/jbc.REV119.011353)

PMID: [31843970](https://pubmed.ncbi.nlm.nih.gov/31843970/)

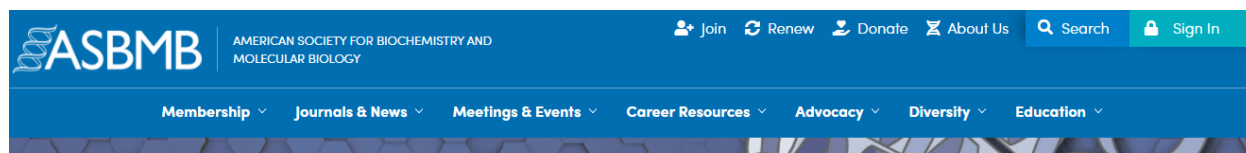
The balancing act of R-loop biology: The good, the bad, and the ugly

[Youssef A. Hegazy](#),[‡] [Chrisan M. Fernando](#),[§] and [Elizabeth J. Tran](#)^{†¶,1}

► [Author information](#) ▾ [Copyright and License information](#) [Disclaimer](#)

[Copyright](#) © 2020 Hegazy et al.

Published under exclusive license by The American Society for Biochemistry and Molecular Biology, Inc.



[Home](#) > [Journals and news](#) > [Copyright and reproduction](#)



Copyright and reproduction

Content in ASBMB journals — the *Journal of Biological Chemistry*, *Molecular & Cellular Proteomics* and the *Journal of Lipid Research* — that was published under those journals' paid open access publishing option, Author's Choice, was distributed under the CC-BY license which automatically grants all commercial and noncommercial use of the article to all, as long as appropriate attribution is given to the original work.

Therefore, permission need not be secured in order to reuse content published under the CC-BY license.

ASBMB journal articles published under the Author's Choice option before April 1, 2013, were initially published under the CC BY-NC license but have been relicensed with the CC-BY license. While readers will continue to see a link on those articles to the CC BY-NC license, **this notice confirms that the license for these articles has been changed to the CC-BY license.**

More details about Author's Choice and other ASBMB journal policies: